EPA 815-R-00-024

GEOMETRIES AND CHARACTERISTICS OF PUBLIC WATER SYSTEMS

December 2000

PREPARED FOR:

U.S. Environmental Protection Agency
Office of Ground Water and Drinking Water
401 M Street, SW
Washington, DC 20460

PREPARED BY:

Science Applications International Corporation (SAIC)
1710 Goodridge Drive
McLean, Virginia 22102

EPA Contract No. 68-C6-0059
SAIC Work Assignment No. 1-24
SAIC Project No. 01-0833-08-3562-000

AND

The Cadmus Group
1901 North Fort Myer Drive, Suite 900
Arlington,VA 22209

EPA Contract No. 68-c-99-206
Cadmus Group Work Assigment No. 1-31

Table of Contents

List of Exhibits

Glossary

Average Daily Flow - daily volume of water produced within a system or by a treatment plant, averaged over 365 days (also called average daily production however).

Community Water System (CWS) - a public water system that has at least 15 service connections used by year-round residents or regularly serves at least 25 year-round residents.

Design Flow - the design capacity, or the maximum amount of water per day that can be treated at the treatment plant or within a system (also called design capacity)

Entry Point - Locations where finished water enters a distribution system or is sold.

Finished Water - water that has passed through a water treatment plant; all the treatment processes are completed or "finished". The water is ready to be delivered to consumers.

Ground Water - water found below the surface of the land, usually in porous rock formations, without significant occurrence of insects, microbes, or pathogens and without rapid shifts in water quality parameters. Ground water is the source of water found in wells and springs.

Ground Water System - water system that gets a majority of its water from ground water.

Ground Water Under the Direct Influence (GWUDI) of Surface Water - any water beneath the surface of the ground with significant occurrence of insects or other microorganisms, algae, or large-diameter pathogens, or significant and relatively rapid shifts in water characteristics such as turbidity, temperature, conductivity, or pH that closely correlate to climatological or surface water conditions. Direct influence must be determined for individual sources in accordance with criteria established by the State.

GWUDI System - water system that gets a majority of its water from GWUDI.

Maximum Daily Flow - the highest flow over one day measured within one year in a system or plant.

Non-Purchased Water System - a system that treats its own water for delivery to the public and does not purchase water from other systems.

Non-Transient Non-Community Water System (NTNCWS) - a public water system that regularly serves at least 25 of the same people more than 6 months per year that is also not a community water system

Public Water System (PWS) - a system for the provision to the public of piped water for human consumption if such a system has at least 15 service connections or regularly serves an average of at least 25 individuals per day at least 60 days out of the year.

Purchased Water System - a system that purchases any amount of drinking water from another system for distribution to its own customers.

Retail Population of a System- population receiving water treated and sold directly to them by that system.

Source Water - The water used as the source for the water treatment plant's operations.

Surface Water - water that is open to the atmosphere and subject to surface runoff.

Surface Water System - water system that gets a majority of its water from surface water (CWSS definition and used throughout the report).

Transient Non-Community Water System (TWS) - a non-community water system that does not regularly serve at least 25 of the same persons more than 6 months per year.

Wholesale Population of a System - the population receiving water treated by a separate water system.

List of Abbreviations

AWWARF - American Water Works Association Research Foundation

BAT - Best Available Technology

CWS - Community Water System

CWSS - Community Water Systems Survey

FRDS - Federal Reporting Data System (now known as SDWIS)

GIS - Geographic Information System

gpcd - gallons per capita per day

gpd - gallons per day

GWUDI - Ground Water Under the Direct Influence of surface water

MWDSC - Metropolitan Water District of Southern California

NCWS - Non-Community Water System

NTNCWS - Non-Transient Non-Community Water System

OGWDW - Office of Ground Water and Drinking Water

PWS - Public Water System

RIA - Regulatory Impact Analysis

SBREFA - Small Business Regulatory Fairness Act

SDWA - Safe Drinking Water Act

SDWIS - Safe Drinking Water Information System

TDP - Technology Design Panel

TWS - Transient Non-Community Water System

UMRA - Unfunded Mandates Reform Act

USGS - U.S. Geological Survey (dept)

1: Introduction

The U.S. Environmental Protection Agency (EPA) historically has analyzed the costs to public water systems (PWSs) and their customers that stem from regulations pursuant to the Safe Drinking Water Act (SDWA). Various regulatory reforms, particularly during the last five years, have also placed considerable emphasis on evaluating the benefits and costs of regulation[1]. Consistent with this trend, Section 103 of the SDWA Amendments of 1996 (codified in Section 1412(b) of the Act) mandates that EPA perform benefit-cost analyses as part of the development process for all new drinking water regulations.

In performing a Regulatory Impact Analysis (RIA) for any drinking water rule under development, EPA must be able to highlight the impacts (i.e., benefits and costs) for typical affected parties, while also capturing the more extreme situations. For example, analysis of new drinking water regulations could address situations that range from campsites in remote national forests to the largest metropolitan areas in the country, such as New York City and Los Angeles. Characterization of data for an RIA must include information on the number of water systems of various types and sizes, average population served, and average and maximum flows in a system. EPA uses these data in various ways to estimate national benefits and costs. For example, costs of a proposed regulation are often estimated by establishing the number of systems of a particular type affected by the rule (usually some proportion of the total systems) and multiplying them by unit costs for implementing additional treatment technologies. To facilitate benefit-cost analyses, system information must be organized into a manageable framework that should inform rather than complicate, while provided adequate precision and accuracy for the necessary evaluations

The purpose of this report is to present a basic set of significant PWS characteristics for use in future Office of Ground Water and Drinking Water (OGWDW) rulemakings. As additional data are gathered and analyzed, the characteristics established in this report can be revised. By describing and encouraging discussion of the underlying data and models used to develop the characteristics, EPA hopes to facilitate acceptance and use of a common set of inputs. A clearer consensus about the basic characteristics of PWSs will help EPA and stakeholders focus attention on RIA results and decision making, rather than on the basic characteristics of PWSs.[2]

The remainder of this document is organized as follows:

Chapter 2: An overview of the universe of PWSs in the United States and a discussion of overarching issues in developing a framework for describing this universe, including characterizing these systems in terms of population served and water source.

Chapter 3: A description of the source and quality of data used to analyze and develop baseline profiles of PWSs.

---

[1]Regulatory benefit-cost analyses on various sectors of the economy are addressed by Executive Order 12866, the Unfunded Mandates Reform Act (UMRA), the Small Business Regulatory Fairness Act (SBREFA), and the Paperwork Reduction Act (PRA).

[2] Contaminant occurrence and technology costs also are important determinants of the accuracy and precision of an RIA. These issues are being addressed as part of other EPA initiatives to improve RIA data and tools.

Chapter 4: EPA's model of community water system (CWS) flow rates, which uses regression equations based on design and average flows and population served.

Chapter 5: EPA's analysis of entry point geometries for CWSs, including distribution of flow among entry points.

Chapter 6: EPA's estimates of numbers and types of drinking water treatment technologies currently in place for CWSs.

Chapter 7: EPA's estimates of the numbers and types of non-community public water systems (NCWs) and typical system sizes (that is, flows and populations served) .

Chapter 8: A list of references used in this document.

2: Description of the Public Water Supply Universe

The universe of PWSs comprises both community and non-community systems (see Exhibit 2.1) with a wide range of characteristics in terms of water flow rates, size and composition of service population, source water types, treatment configurations, types of ownership, etc. As mentioned in Chapter 1, regulatory analysis requires a manageable framework for describing this universe. This chapter describes EPA's existing model as well as the rationale for revising this model. An overview of the PWS universe is then presented, including an inventory of the number of systems by population category, water source, and ownership type. Finally, overarching issues are considered, including issues related to the measurement of population served and source water type.

Exhibit 2.1
40 CFR §141.2 Definitions

Public water system (or PWS) means a system for the provision to the public of piped water for human consumption, if such system has at least 15 service connections or regularly serves an average of at least 25 individuals daily at least 60 days out of the year. Such term includes (1) any collection, treatment, storage, and distribution facilities under control of the operator of such system and used primarily in connection with such system, and (2) any collection or pretreatment storage facilities not under such control which are used primarily in connection with such system. A PWS is either a ''community water system'' or a ''noncommunity water system.''

Community water system **(or *CWS*)** means a public water system that serves at least 15 service connections used by year-round residents or regularly serves at least 25 year-round residents.

Non-community water system **(or *NCWS*)** means a public water system that is not a community water system.

Non-transient non-community water system **(or *NTNCWS*)** means a public water system that is not a community water system and that regularly serves at least 25 of the same persons more than 6 months per year.

Transient non-community water system **(or *TWS*)** means a non-community water system that does not regularly serve at least 25 of the same persons more than 6 months per year.

2.1 Existing Model

The existing model stratifies the drinking water universe into 12 population size categories. For each population size category, the median population defines a typical drinking water system's size. Average daily flow (volume of water produced per day) and design flow (design capacity), expressed in million gallons per day for the corresponding population, are also included in the model. Exhibit 2.2 depicts the existing model for community water systems and its 12 strata, median populations, and flows.

Exhibit 2.2
Existing Model of Community Water System Characteristics

Population Category	Median Population	Average Flow (million gallons per day)	Design Flow (million gallons per day)
25 to 100	57	0.0056	0.024
101 to 500	225	0.024	0.087
501 to 1,000	750	0.089	0.27
1,001 to 3,300	1,910	0.23	0.65
3,301 to 10,000	5,500	0.70	1.8
10,001 to 25,000	15,500	2.7	4.8
25,001 to 50,000	35,000	5.0	11.0
50,001 to 75,000	60,000	8.8	18.0
75,001 to 100,000	88,100	13.0	26.0
100,001 to 500,000	175,000	27.0	51.0
500,001 to 1,000,000	730,000	120.0	210.0
Greater than 1,000,000	1,550,000	270.0	430.0

The information in Exhibit 2.2 was used by regulatory analysts to estimate the cost for a typical system for each of the 12 population size categories. Capital and operation and maintenance costs for the typical system were based on design and average flows, respectively. Costs for the typical system were then extrapolated to the national level for each population size category based on a relatively simple probability decision tree. The sum of the costs for each population category yielded the national compliance cost for a given regulatory scenario.

With the exception of system counts, a comparable model describing the flow and population characteristics does not exist for non-community water systems.

2.2 Rationale for Revising the Existing Model

One of EPA's objectives is to expand the range of the existing model for drinking water treatment profiles. The impetus for revising the existing model stems from Agency experience in promulgating drinking water treatment regulations during the 1980s and early 1990s. EPA's OGWDW recognizes the need to improve existing tools and models for regulatory impact analysis to more adequately describe regulatory impacts. The need to develop improved methodologies came into focus with the advent of a new regulatory climate in the 1990s; the passage of the Unfunded Mandates Reform Act (UMRA), the Small Business Regulatory Fairness Act (SBREFA), and Executive Order 12866; the reauthorization of the SDWA in 1996; and calls from stakeholders to improve process used to estimate benefits and costs for upcoming drinking water regulations.

In 1996, EPA convened a Blue Ribbon Panel (the Panel) of experts from academia, the water treatment industry, and State, local, and Federal governments to help critically evaluate various components of the Agency's regulatory analysis approach. One of the components addressed by the Panel was the issue of modeling PWS characteristics. Public comments and published reports suggested that EPA's existing water system profile needed to be revised to accommodate more sophisticated analysis and to improve their capability to assess impacts for various types of public water systems.

The Panel's recommendations that apply to improving the water industry profiles follow:

- Examine the relationship between population served and flow. Examine whether population truly relates to flow in a system. The Panel noted that the relationship does hold up on average but variations across systems in per capita flow (as much as 10 to 1) can be expected.
- Investigate the design-to-average flow ratios for small systems. The Panel believes that these ratios may be high. For medium to large systems, a reasonable ratio is between 1 and 2.
- EPA should incorporate greater diversity into the analysis so the results support a variety of objectives and inquiries. The Panel proposed a classification scheme based on three primary variables: system size, type of ownership, and source water.
- System size categories should account for the differences in technical, financial, and managerial characteristics. At least five general categories are necessary to capture this diversity. Very small, small, medium, large, and very large systems should be captured in the scheme.
- EPA should devote additional analysis to systems that have more than one treatment plant or "entry point" into the distribution system.

In addition to strictly flow-based models, the Panel also suggested looking at other variables to develop profiles. This work is being carried out under a separate assignment.

Recommendations by the American Water Works Association Research Foundation (AWWARF) also focused on developing profiles based on source water characteristics. AWWARF concluded that average daily and design flows are different for ground and surface water systems, and that the number of entry points into a system affect compliance costs.

The Panel's and AWWARF's recommendations provided basic guidelines for developing revised models to characterize PWSs. The existing model described above does not have the capacity necessary to address some of the recommendations. For example, the data set on which the existing model is based does not distinguish systems by ownership or by source water category, nor did it address characteristics such as numbers of entry points per system. To address the various concerns and develop a revised model of drinking water treatment profiles, EPA turned to the two most comprehensive sources of information available on the full spectrum of drinking water systems: EPA's Safe Drinking Water Information System (SDWIS), described below, and the Community Water Systems Survey (CWSS) of 1995, described in Chapter 3.

2.3 Characterization of the Water Supply Universe

SDWIS provides the most complete inventory of the U. S. water supply industry. It contains information about public water systems and their violations of EPA regulations for safe drinking water. The inventory is used not only for compliance tracking purposes, but also to assist in allocating grant monies among the States. A considerable effort is expended to ensure that SDWIS accurately fulfills these needs.[3]

Near the end of each calendar year, a snapshot of the SDWIS inventory is distributed to State drinking water programs for verification of the number and types of systems, a process that customarily takes several months. The inventory reflected in this report were derived from the December 1998 database. Because both population figures and system counts change continuously, these figures should be considered representative of a particular time, not a static universe. For example, those performing risk analysis should consider that the number of private systems has been increasing over the years (Dysard, 1999), and some believe that the number of wholesale systems will increase as well. There has also been a steady increase in per capita water use, which will affect system average and design flow data (Linsley et. al, 1992). Notwithstanding, the figures presented in this report represent the broad universe of populations and systems to be considered in the risk assessment. Information contained in SDWIS is complete for the categories identified by the Blue Ribbon Panel as important for industry subcategorization. Core verified data in the inventory include:

- Federal identification (ID) number
- Source water (ground, surface, and ground water under direct influence of surface water (GWUDI)
- Ownership type (local government, private, mixed, etc.)
- Regulatory classification (community, transient, etc.)

Further discussion of the characteristics of these data elements and how they are proposed to be used in regulatory impact analyses is provided in Sections 2.4 and 2.5. Exhibits 2.3 through 2.5 present the numbers of systems subcategorized into traditional analytical categories. The ownership classification is limited to private versus public since this the only distinction shown to have appreciable effects on the technology cost models (Chapter 4 addresses this issue). Systems are further categorized by source water (surface, ground, or GWUDI) and by whether they are purchased or non-purchased systems. Population figures for the various sizes of water systems are presented in Appendix A**.**

2.4 Population Served

PWSs serve commercial, industrial, and residential customers. While it is generally true that systems distribute their treated water directly to their customers, there are cases where water is wholesaled to another utility that subsequently distributes water to customers. Thus, commercial, industrial, and residential customers can be part of the retail population of a system (i.e., they receive water directly from that system) or they can be part of the wholesale population of a system (i.e., they receive water from a second system that buys their water from the first system). Exhibit 2.6 shows the link between source water, treatment plants, and residential and wholesale population served by a system. One of the

[3] It is important to note, however, that complete names and addresses are not available in some States. The absence of this core information suggests the inventory may still include inactive systems.

Exhibit 2.3 Community Water Systems (Number of Systems)

Primary Source/ Ownership	Population Category										
	Less than 25	25 to 100	101 to 500	501 to 1,000	1,001 to 3,300	3,301 to 10,000	10,001 to 50,000	50,001 to 100,000	100,001 to 1,000,000	Greater than 1,000,000	Total
Ground Water	178	14,099	15,058	4,689	5,714	2,459	1,216	131	61	2	43,607
Public	12	1,194	4,209	2,608	3,888	1,913	979	104	46	2	14,955
Private	155	12,244	9,694	1,636	1,428	435	201	26	11	0	25,830
Purchased-public	0	129	469	260	279	80	32	1	4	0	1,254
Purchased-private	3	156	304	106	76	15	3	0	0	0	663
Other	8	376	382	79	43	16	1	0	0	0	905
Surface Water	75	921	1,899	1,178	2,356	1,802	1,592	299	259	13	10,394
Public	18	142	382	342	938	899	829	150	169	12	3,881
Private	9	216	303	84	143	62	93	29	38	1	978
Purchased-public	40	181	697	546	1,014	719	593	109	46	0	3,945
Purchased-private	8	267	437	174	220	105	68	10	6	0	1,295
Other	0	115	80	32	41	17	9	1	0	0	295
GW Under SW Influence	0	96	104	42	64	42	15	1	2	0	366
Public	0	13	31	24	48	35	12	1	2	0	166
Private	0	73	59	9	3	4	0	0	0	0	148
Purchased-public	0	0	7	5	12	3	2	0	0	0	29
Purchased-private	0	2	0	1	0	0	0	0	0	0	3
Other	0	8	7	3	1	0	1	0	0	0	20
Total	253	15,116	17,061	5,909	8,134	4,303	2,823	431	322	15	54,367
Public	70	1,659	5,795	3,785	6,179	3,649	2,447	365	267	14	24,230
Private	175	12,958	10,797	2,010	1,870	621	365	65	55	1	28,917
Other	8	499	469	114	85	33	11	1	0	0	1,220

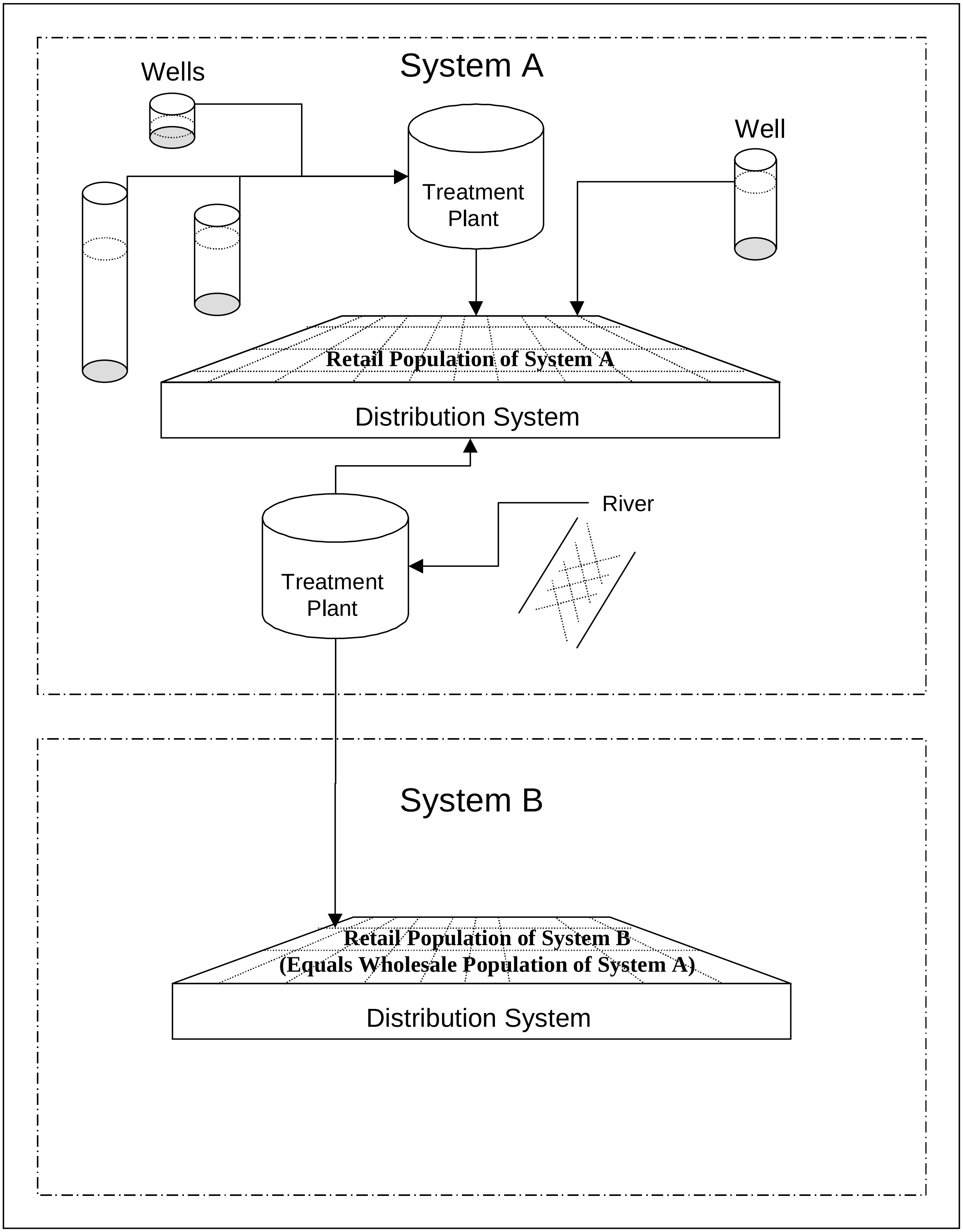
Exhibit 2.4 Nontransient Non-Community Water Systems (Number of Systems)

Primary Source/ Ownership	Population Category										
	Less than 25	25 to 100	101 to 500	501 to 1,000	1,001 to 3,300	3,301 to 10,000	10,001 to 50,000	50,001 to 100,000	100,001 to 1,000,000	Greater than 1,000,000	Total
Ground Water	31	9,659	7,038	1,978	692	62	14	0	0	0	19,474
Public	4	1,711	3,097	1,156	332	19	6	0	0	0	6,325
Private	24	7,440	3,714	771	339	38	4	0	0	0	12,330
Purchased-public	0	10	11	7	5	4	3	0	0	0	40
Purchased-private	0	22	17	6	0	0	1	0	0	0	46
Other	3	476	199	38	16	1	0	0	0	0	733
Surface Water	3	258	284	102	79	22	4	1	1	0	754
Public	2	42	33	9	18	3	0	0	0	0	107
Private	0	71	122	49	34	8	0	0	0	0	284
Purchased-public	0	13	26	6	6	3	3	1	1	0	59
Purchased-private	0	71	39	12	10	4	1	0	0	0	137
Other	1	61	64	26	11	4	0	0	0	0	167
GW Under SW Influence	1	9	11	4	1	1	0	0	0	0	27
Public	0	1	4	4	0	0	0	0	0	0	9
Private	1	8	7	0	0	1	0	0	0	0	17
Purchased-public	0	0	0	0	0	0	0	0	0	0	0
Purchased-private	0	0	0	0	0	0	0	0	0	0	0
Other	0	0	0	0	1	0	0	0	0	0	1
Total	35	9,926	7,333	2,084	772	85	18	1	1	0	20,255
Public	6	1,777	3,171	1,182	361	29	12	1	1	0	6,540
Private	25	7,612	3,899	838	383	51	6	0	0	0	12,814
Other	4	537	263	64	28	5	0	0	0	0	901

Exhibit 2.5 Transient Non-Community Water Systems (Number of Systems)

Primary Source/ Ownership	Population Category: Less than 25	25 to 100	101 to 500	501 to 1,000	1,001 to 3,300	3,301 to 10,000	10,001 to 50,000	50,001 to 100,000	100,001 to 1,000,000	Greater than 1,000,000	Total
Ground Water	660	71,107	19,137	1,925	616	150	71	12	9	0	93,687
Public	60	8,865	2,585	602	261	81	38	7	8	0	12,507
Private	587	56,723	15,793	1,234	308	54	11	5	1	0	74,716
Purchased-public	1	72	39	13	17	2	0	0	0	0	144
Purchased-private	5	550	49	11	2	1	0	0	0	0	618
Other	7	4,897	671	65	28	12	22	0	0	0	5,702
Surface Water	33	1,275	491	85	58	28	7	3	0	0	1,980
Public	2	165	129	30	29	19	4	1	0	0	379
Private	21	493	223	27	12	3	0	0	0	0	779
Purchased-public	6	86	26	10	10	2	1	0	0	0	141
Purchased-private	3	351	62	13	4	1	0	0	0	0	434
Other	1	180	51	5	3	3	2	2	0	0	247
GW Under SW Influence	2	53	26	3	3	0	0	0	0	0	87
Public	0	14	9	3	0	0	0	0	0	0	26
Private	2	39	14	0	2	0	0	0	0	0	57
Purchased-public	0	0	2	0	0	0	0	0	0	0	2
Purchased-private	0	0	0	0	0	0	0	0	0	0	0
Other	0	0	1	0	1	0	0	0	0	0	2
Total	695	72,435	19,654	2,013	677	178	78	15	9	0	95,754
Public	69	9,202	2,790	658	317	104	43	8	8	0	13,199
Private	618	58,156	16,141	1,285	328	59	11	5	1	0	76,604
Other	8	5,077	723	70	32	15	24	2	0	0	5,951

Exhibit 2.6
Representation of System Relationships



most complicated examples of these relationships is the Metropolitan Water District of Southern California (MWDSC), which has been estimated to serve between 8 and 16 million people, all through wholesale relationships (purchased water systems).

As noted earlier, the primary source of information on water treatment utilities is SDWIS, which tracks population served by a system on a retail customer basis only. Comprehensive information on total (retail plus wholesale) populations does not exist except for the largest water systems. This data reporting approach is useful for many reasons, but, as can be seen by the MWDSC example above, any analysis using these data must correctly interpret the results and explore the potential or degree of bias.

The current system categorization scheme within SDWIS introduces a bias into the cost analysis. Since water systems are classified based on retail population served, systems that purchase their water from another system and distribute it are considered stand-alone systems serving only their retail customers. The overall result is that this classification process accounts for the total national retail population and flows, but assumes more individual systems of smaller sizes treating their water than actually occurs. Costs are then higher than would actually occur because no economies of scale are available (i.e., it is less expensive on a unit cost basis to install a technology at larger plant than to install technologies at smaller plants).

The next several paragraphs summarize an attempt to quantify the cost bias. The retail population served by purchased water systems were allocated to nonpurchased systems and then theoretical costs were generated using the two classification schemes. Detailed descriptions of the analyses and all associated assumptions and calculations are included in Appendix B.

Step 1: Evaluate January, 2000 SDWIS data to determine the degree to which CWSs that purchase water do so from systems of similar sizes or of different sizes. This analysis addressed only cases where primary source of the buyer and seller is the same and did not include systems with cascading provider relationships (i.e., where a seller provides water to a purchased system which in turn sells water to another purchased system). The analysis evaluated surface water, ground water, and GWUDI systems separately and looked only at four major size categories of systems (very smalls, smalls, mediums, and larges). GWUDI systems are not included in subsequent analyses because they represent such a small portion of the public water system universe.

Step 2: Allocate the populations for the purchased water systems to nonpurchased water systems to estimate the effective shift to higher size categories. This was done by estimating the mean population for the purchased water system and allocating it to the nonpurchased systems based on the percentages developed in Step 1. The smallest impacts were observed when very small systems buy from any size category (the new median population is not large enough to move a system into another size category), and the largest impacts are seen when the large systems buy from any size category (no matter the size of the seller, they all become large systems). Exhibits 2.7 and 2.8 summarize the results of this step for surface and ground water systems, respectively. The first three rows summarize the total CWS universe on a retail population basis. The final row presents the number of systems as they might exist if retail and wholesale population were combined for modeling purposes.

Exhibit 2.7
Comparison of System Categorization Schemes for Surface Water Systems

System Type	Number of Surface Water CWSs by System Size Category*									
	25 to 100	**101 to 500**	**501 to 1,000**	**1,001 to 3,300**	**3,301 to 10,000**	**10,001 to 50,000**	**50,001 to 100,000**	**100,001 to 1,000,000**	**Greater than 1,000,000**	**Totals**
Nonpurchased Systems (a)	382	696	411	1,086	958	913	178	200	14	4,838
Purchased Systems (b)	483	1,185	719	1,282	828	603	89	47	0	5,236
Total Systems , Retail Based (a)+(b)	**865**	**1,881**	**1,130**	**2,368**	**1,786**	**1,516**	**267**	**247**	**14**	**10,074**
Total Systems, Retail + Wholesale Based	**198**	**362**	**461**	**1,217**	**1,074**	**1,080**	**210**	**221**	**16**	**4,838**

*Nonpurchased and purchased systems represent the sum of public and private systems for this analysis. Does not include "other" ownership category.

Exhibit 2.8
Comparison of System Categorization Schemes for Ground Water Systems

System Type	Number of Ground Water CWSs by System Size Category*									
	25 to 100	**101 to 500**	**501 to 1,000**	**1,001 to 3,300**	**3,301 to 10,000**	**10,001 to 50,000**	**50,001 to 100,000**	**100,001 to 1,000,000**	**Greater than 1,000,000**	**Totals**
Nonpurchased Systems (a)	13,438	13,903	4,244	5,316	2,348	1,180	130	57	2	40,618
Purchased Systems (b)	285	773	366	355	95	35	1	4	0	1,914
Total Systems , Retail Based (a)+(b)	**13,723**	**14,676**	**4,610**	**5,671**	**2,443**	**1,215**	**131**	**61**	**2**	**42,532**
Revised Systems, Retail + Wholesale Based	**13,401**	**13,864**	**4,270**	**5,349**	**2,362**	**1,181**	**130**	**59**	**2**	**40,618**

*Nonpurchased and purchased systems represent the sum of public and private systems for this analysis. Does not include "other" ownership category.

Step 3: Estimate national costs of treatment using the two categorization schemes. Three technologies were selected from the EPA Document, "Technologies and Costs for Control or Microbial Contaminents and Disinfection Byproducts" (November 2000). Total annual cost per system (yearly operation and maintenance costs and annualized capital cost) were estimated for each size category, and national costs were calculated using the two system categorization schemes (retail based and retail+wholesale based).

Results from step three show that for surface water systems, national costs of treatment could be 22 to 45 percent higher than what might be incurred if retail+wholesale based system categorization is used. The effects on cost estimates for ground water systems is much less (approximately 4 percent increase). Despite this bias, EPA believes that the certainty afforded benefits estimates through the use of the SDWIS inventory justifies its use.

2.5 Source Water

Another area of interest to the Blue Ribbon Panel and others relates to distinguishing systems on a source water basis. While most small water systems have one source only, it is not unusual for larger water systems to have multiple sources of water supply. These source waters may include ground water, surface water, and ground water under the direct influence of surface water (GWUDI). Most systems, however, have predominately one source water type. Many past regulatory estimates have not modeled mixed systems separately. To the extent that occurrence profiles differ in ground versus surface water, or when regulations only impact one type, accounting for the numbers of these mixed systems is important.

SDWIS defines any water system with a continuous input of surface water as a surface water system. This is the case even if 99 percent of the water is of ground water origin. The CWSS, however, distinguishes and groups systems based on the predominant source water type, or the source water(s) that supply more than 50 percent of water for the entire system. The CWSS not only categorizes the entire system by source water type, but further categorizes each entry point into the distribution system by source water type and treatment.

Extrapolating from the CWSS information, Exhibit 2.8 presents an estimate of the number of non-purchased mixed systems that SDWIS classifies as surface water systems. Exhibit 2.8 also summarizes the number of systems out of the total inventory that would require regrouping if SDWIS had used predominant source type as the classification scheme. As shown in the exhibit, it is estimated that 1,069, or 21 percent, of non-purchased surface water systems in SDWIS have some ground water source flow, and 435, or 8 percent, of the non-purchased surface water systems get the majority of flow from ground water sources. The CWSS did not provide enough information to perform a similar analysis for purchased water systems. Analysts performing inventory subcategorizations by source need to carefully consider these numbers when performing regulatory impact analyses.

Exhibit 2.9
Analysis of Mixed Systems in SDWIS Non-Purchased Surface Water Systems

System Type		Population Category									
		Less than 100	**100 to 500**	**501 to 1,000**	**1,001 to 3,300**	**3,301 to 10,000**	**10,001 to 50,000**	**50,001 to 100,000**	**100,001 to 1,000,000**	**Greater than 1,000,000**	**Total**
(1)	Total Ground Water Systems (SDWIS)	14,391	15,070	4,739	5,726	2,489	1,282	139	70	2	43,908
(2)	Total Non-Purchased Surface Water Systems (SDWIS)	599	853	473	1,179	1,008	934	180	200	14	5,440
(3)	Number (and %) of Surface Water Systems with Ground Water Component*	22 (3.7%)	123 (14%)	33 (7.0%)	225 (19%)	264 (26%)	249 (27%)	68 (38%)	80 (40%)	5 (35.7)	1,069 **(21%)**
(4)	Number (and %) of Surface Water Systems with Majority Flow from Ground Water Sources*	22 (3.7%)	82 (9.6%)	0	69 (5.9%)	125 (12%)	93 (10%)	16 (8.9%)	28 (14%)	0	435 **(8.4%)**

* Extrapolated from CWSS Data

3: Community Water System Survey of 1995

As discussed in Chapter 2, EPA used two primary data sources in this study: SDWIS and the 1995 CWSS. SDWIS was used to develop the detailed inventory of public water systems presented in Chapter 2. The CWSS provided the data necessary to analyze the characteristics of public water systems: population and flow relationships, number and types of treatments-in-place, and entry points into the distribution system. The CWSS is described in detail below, along with the reasoning behind its selection as a primary data source.

The EPA OGWDW periodically conducts surveys of the financial and operating characteristics of community water systems. The most recent of these is the 1995 CWSS. The purpose of the 1995 CWSS was to collect information that would do the following:

- Help EPA and States develop and implement proposals for reauthorizing the Safe Drinking Water Act (SDWA)
- Facilitate water system capacity/development
- Help determine the need for and design of Best Available Technology (BAT) programs
- Support economic and financial analyses of new and revised regulations
- Help EPA identify, evaluate, and provide guidance for best management practices

Because the purpose of the CWSS is closely aligned with the purpose of this analysis, the survey provides an excellent source of information for subsequent chapters of this report. The following specific characteristics of the CWSS make it a useful data source:

- The survey collected the type of data required for this analysis (e.g., population, flow, treatment technologies)
- The sampling method allowed generally accepted statistical protocols
- The survey was specifically designed to capture systems with differing sizes, types of ownerships, and water sources
- The survey incorporated extensive peer review of its design and the results were subjected to extensive quality-assurance procedures
- Specific validation efforts focused on the data critical to this analysis (specifically, population, flow, and treatment facility information)
- The survey data are readily available and amenable to the additional screening required for this analysis

The sections below provide additional background on the CWSS, its statistical design, and the quality-assurance efforts incorporated in the survey. These sections provide further details on the characteristics discussed above and illustrate the survey's usefulness for this analysis. Much of the discussion below is based on information provided in the EPA report "Community Water System Survey: Volume II: Detailed Survey Result Tables and Methodology" (January 1997).

3.1 Survey Overview

EPA began the 1995 CWSS in the fall of 1994. The survey included two phases: a telephone screening survey (Phase I) and a substantive mail survey (Phase II). Phase I was targeted toward a sample of 5,856 water systems from the more than 57,000 systems identified in the Federal Reporting Data System (FRDS), EPA's registry of public water systems (now known as SDWIS). The purpose of Phase I was to identify water systems eligible for the Phase II mail survey and the appropriate contacts for the Phase II survey. Phase I also collected basic data on system size, ownership, and water sources to verify the preliminary information from FRDS and contribute to the design (i.e., set the "sampling frame") of Phase II. In the Phase I screening, conducted from November to December 1994, 4,729 eligible community water systems were identified.

Based on the Phase I findings, 3,681 systems were selected to receive the Phase II mail survey. Phase II involved collecting a variety of substantive operating and financial data, including information about the following:

- Production and storage
- Distribution
- Operator training
- Water sources and treatment
- Source water protection
- Revenues and expenses
- Assets, liabilities, and debt
- Capital investment

Of particular importance for this analysis, information on population served, drinking water flow rates (average daily flow, peak daily flow, and maximum daily treatment design flow), treatment systems in place, and number and location of entry points into the distribution system was collected during Phase II. Appendix C presents the questionnaire used in Phase II. The mail survey was conducted from June 1995 through March 1996. A total of 1,980 systems (approximately 50 percent of those surveyed) responded. Although not every system responded to every question, the majority of systems provided the data crucial to this analysis (e.g., population, flow). Thus, the CWSS provides a substantial database for use in this analysis.

3.2 Statistical Design

This section describes those elements of the CWSS's statistical design that are significant for the purposes of this analysis. The EPA report, "Community Water System Survey: Volume II: Detailed Survey Result Tables and Methodology," contains a more complete and detailed description of the survey design.

To ensure that the results would capture a range of system sizes, types of ownership, and water sources, the CWSS utilized a stratified sample design. A stratified sample is appropriate when subpopulations within the larger population are expected to differ from one another in meaningful ways. In stratified sampling, the population is first divided into subpopulations called strata and random samples are selected from within each stratum. This technique ensures adequate statistical representation for each subpopulation.

Phase I of the CWSS defined strata based on eight categories of size, two categories of ownership, and two categories of source water, for a total of 32 strata. Responses to Phase I resulted in identifying inaccuracies in the initial placement of systems into strata. That is, the Phase I responses showed that the size, ownership, or source water categorization of some systems was different from that expected based on the initial FRDS data. This reclassification of systems, or "stratum migration," required the use of a more complex stratification scheme in Phase II in order to obtain optimum sampling rates; 37 strata were used during Phase II.

Interpreting stratified sampling results at the full population level requires the use of sample weights. The importance of an individual system's survey response depends on how much of the stratum it represents. For example, if one samples 100 systems out of a stratum with a total population of 200 systems, the base weight of each sampled system is 200/100 = 2. These base weights must be adjusted to account for nonresponses during sampling. When some systems in a stratum do not respond, the proportion of the stratum represented by each respondent changes.

For the CWSS, base weights were adjusted weights for nonresponses in both Phase I and Phase II. Further adjustments were made to account for the new strata introduced by stratum migration after Phase I. Weights were adjusted to account for aggregation in the responses (i.e., some respondents submitted combined responses for multiple systems). Finally, weights were "trimmed" for some systems, with extreme weights to reduce variation and increase the precision of sampling estimates. The EPA report, "Community Water System Survey: Volume II: Detailed Survey Result Tables and Methodology," describes each of these adjustments. All of these adjustments resulted in a final weight for each survey response, which was reported along with the survey results.[4]

In addition to characterizing a stratified population, another of the survey's design objectives was to achieve a minimum statistical confidence level. Specifically, the number of samples taken from within each stratum had to be sufficient to obtain estimates with an error not exceeding 10 percent at the 95 percent confidence level. That is, if 50 percent of sampled systems in a stratum reported a certain characteristic, EPA could be 95 percent confident that between 40 and 60 percent of the full population of the stratum have that characteristic. Because of stratum migration and nonresponses, the CWSS did not quite achieve this confidence level for all strata. However, most strata did meet this confidence level and the maximum error did not exceed ±15 percent for any stratum.

One additional relevant characteristic of the CWSS design is the actual sampling strategy employed. Within each stratum, candidate systems were sorted by EPA region and, within each region, by population served. The CWSS then used systematic equal probability sampling to select the surveyed systems. This approach ensured geographic representation of the systems sampled and increased the probability that a range of population sizes within a stratum was represented.

All these elements of the survey's design are relevant for the purposes of this analysis. The sampling design allows characterization of systems with different sizes, ownerships, and source water types, while

[4] The final weights were further adjusted for item-level nonresponse. The process and rationale for making these adjustments is discussed in Chapter 4.

ensuring geographic representation. The sampling weights facilitate the use of data in modeling. The achievement of reasonable reliability imparts confidence in estimates based on CWSS data.

3.3 Peer Review and Quality Assurance

Prior to its implementation, the CWSS was the subject of peer review and testing. Draft versions of the survey questionnaire were peer reviewed by representatives of the National Rural Water Association, the American Water Works Association, and the National Drinking Water Advisory Council; by a consultant from the Government Finance Group; and by an independent consultant specializing in the operational characteristics of drinking water systems. The questionnaire was pretested with nine water systems in Maryland and Delaware. Following the pretest, the full survey process (the sampling routine, Phase I telephone screening, and Phase II mail survey) was pilot tested with 81 systems. As a result of this review and testing, the survey designers made improvements to the sampling plan and to telephone interviewer training. There also were changes in the terminology, content, and structure of both the Phase I and Phase II surveys. These changes increased the likelihood that respondents correctly interpreted the survey questions which enhanced the validity of the results.

The CWSS also incorporated extensive quality assurance (QA) procedures during and after both Phase I and Phase II. QA during Phase I included automated online response checks and periodic staff review of accumulated results. After completion of Phase I, the results were reviewed against FRDS data. QA during Phase II included supervision and spot checking during mailing preparation, pre-data-entry editing, independent double-key entry to minimize data entry errors, and automated data range checks during entry. At the end of Phase II, a final automated data validation effort included statistical evaluation, cross-tabulation checks of related variables, and internal logic checks. The purpose of this validation was to verify consistency and reasonableness and to guide expert review of individual responses. The automated validation examined 500 of the 600 survey variables, including all those used in this analysis. Problems resolved by this process included order-of-magnitude reporting errors, such as the use of gallons instead of million gallons for some questions. At the conclusion of the validation process, a few extreme data outliers (approximately one quarter of 1 percent of the data points) were excluded from the results.

In addition to the in-process QA and automated validation, the CWSS also included manual validation and expert review of responses to eight critical survey questions. These included several questions of importance to this analysis: sources of water (including the data from which average daily flows were derived), population served, and treatment facility information. The manual validation process was used to review answers to these questions for completeness and internal consistency with answers to other questions. When problems were found, the reviewers attempted to derive an answer using responses to other questions, estimate an answer using best professional judgment, or contact the respondent for clarification. Examples of corrected problems include incorrect units and mathematical errors.

The extensive review, testing, and QA incorporated in the CWSS allow increased confidence in the validity of the survey results. The additional manual validation of the data elements used in this study provides further assurance of a realistic basis for further analysis.

4: Analysis of Population and Flow Relationships

This chapter presents a model for defining CWS size categories for regulatory analysis. The model is based on a regression analysis of the relationship between flow and population. Section 4.1 presents the rationale for selecting flow and population as the variables for performing the regression. Sections 4.2 and 4.3 present the analysis of average daily flow and maximum daily treatment capacity (design flow) for CWSs.

4.1 Population and Flow as Critical Variables

Population and flow (average daily and design), are two key variables in the development of regulatory impact analyses. A system's average daily and design flows are driving factors in estimating potential operating and capital expenditures. The corresponding population served identifies the number of people who will derive benefits from compliance. Further, population served, when coupled with exposure information, provides a basis for estimating household benefits and costs of regulatory alternatives.

EPA studies[5] and the published literature point to population and flow variables as basic defining features of CWSs. Thus, an analysis of population and flow is a logical starting point for modeling the characteristics of drinking water systems for regulatory benefit and cost analysis. There are more sophisticated means of forecasting urban water use than basic flow and population equations; however, considering the water characteristics of interest described above, the single variable approach for predicting flow based on population will meet regulatory analysis needs. Also, other variables do not readily lend themselves to developing a model for estimating regulatory benefits and costs. For example, variables linked to a system's water sales could be used, but revenue information is more difficult to obtain and not as easily linked to the amount of water to be treated.

Acceptance of population and flow as critical variables still leaves an important issue to be resolved. As discussed in Chapter 2, EPA tracks systems based on retail population served. This criterion also was used in the stratification of the CWSS (i.e., respondents were asked to report the number of people served by their systems and number of residential connections). Treatment expenditures, however, relate to total water treated. Consequently, if EPA used retail population with total flow estimates, double counting would result. In particular, the wholesale portion of flow would be repeat-counted when costs are estimated for purchased water systems. While one could eliminate purchased systems from the cost/benefit analysis to avoid this double counting, data are not available on wholesale customers served. Even if double counting were avoided, inclusion of wholesale flows without including the associated population would bias the flow model by distorting the population/flow relationship. Household costs would be overestimated.

For all of the aforementioned reasons, this analysis adjusts system flows to remove sales for resale from wholesale systems prior to regression analysis. The net effect of this approach is to "assign" these flows to the purchased water systems. This assignment makes it possible to use the SDWIS inventory directly for compliance cost estimation. The disadvantages of this modeling approach are twofold:

---

[5] Cummins, Michael D. 1987. *Analysis of Flow Data*. Report prepared for EPA, Office of Drinking Water. October 5.

- Costs and benefits to the largest systems may be underestimated. Because half of the 65 largest utilities wholesale to smaller utilities, the associated costs and benefits will appear in lower size categories.

- Total national costs will be overestimated. Due to economies of scale, unit treatment costs generally vary inversely to system size. Wholesale flows from larger systems that reflect economies of scale are assigned to smaller purchased water systems that otherwise do not capture these economies of scale.

The impact of the first disadvantage will be addressed by not applying these models to the largest systems; rather, EPA will attempt specific estimation based on actual configurations for these systems. As for the second disadvantage, the affect of this bias has been quantified in Chapter 2.

4.2 Analysis of Average Daily Flow versus Population

A regression analysis of average daily flow and a system's corresponding retail population was performed. A regression analysis examines the nature and strength of the relationship between a variable of interest (known as a dependent variable) and one or more other variables (known as independent or explanatory variables) that are believed to affect the first variable. In this case, average daily flow is the dependent variable and retail population is the single independent variable.

Based on previous studies, a water system's flow is known *a priori* to be strongly dependent on population served. Therefore, the purpose of this regression analysis is to (1) confirm the strength of this relationship using the statistically sound CWSS data set, and (2) develop a model or models describing the relationship between population and flow.[6]

4.2.1 Data Screening for Regression Analysis

Prior to regression analysis, the CWSS database of 1,980 respondents was screened for missing data points. The initial screening procedure entailed a two-step process: applying a formula for average daily flow followed by the elimination of nonresponses and zero values. In the first step, adjustments were made to average daily flow reported in Question 4 of the CWSS questionnaire (see Appendix C). Question 4 reports total annual surface water, ground water, and purchased water flow.

The reported flow in Question 4 represents the total flow of a system. This may include wholesale flow, which is the portion of total average daily flow the system sells to other water utilities. To obtain retail flow, total flow reported in Question 4 was adjusted by subtracting wholesale flow as described below.

---

[6] In the CWSS, flow is reported at the system-level and at the entry point level. An entry point is typically a treatment plant but can also be any location where potential treatment could be installed. Systems (particularly large ones) often can have more than one entry point. Analysis of flow and population relationships can be performed at either the system or entry point level. Since population is reported in the CWSS only at the system level, the regression analysis was performed at the system level. Analysis of entry point characteristics are presented in Chapter 5.

Question 29 of the CWSS questionnaire reports a facility's total production flow divided into various uses (see Appendix C). These uses include water delivered to residential, commercial, industrial, wholesale, local, and other customers. In theory, the flow reported in Question 4 should match the sum of the flows reported in Questions 29 and 32 (losses and water supplied free to municipal uses). Where significant discrepancies existed, responses were omitted. As flows had been subjected to the highest level of QA in conducting the survey, it was believed that other categories would be, at best, no more reliable.

The following formula was used to adjust the dependent variable (i.e., average daily flow):

$$\textit{Average Daily Flow} = \frac{(\textit{Question 4 Annual Flow} - \textit{Question 29 Annual Wholesale Flow})}{365\ \textit{days}}$$

This formula was applied to all 1,980 respondents in the CWSS database. Following the adoption of this convention, the data in Questions 4 and 11 were screened. Of the 1,980 respondents, 184 systems did not report adequate flow information in Question 4, and 44 did not report a population or the number of service connections in Question 11. Nine of the largest systems also were eliminated from the data set since these systems will be analyzed on a site-specific basis.[7] An additional nine systems were eliminated based on either population or flow-reporting discrepancies. Ninety-nine systems did not report a population but reported the number of service connections. Populations were imputed for these systems by using a population-served to number of service connection ratio. For the eight population categories in CWSS, population served to number of service connection ratios shown in Exhibit 4.1 were used to make this adjustment.

Exhibit 4.1. Population-Served to Number of Service Connections Ratios

System Size (Population Served)*	Less than 100	101 to 500	501 to 1,000	1,001 to 3,300	3,301 to 10,000	10,001 to 50,000	50,001 to 100,000	100,001 to 500,000
Ratio	2.3	2.4	2.6	2.9	3.0	3.7	3.8	5.3

* CWSS population categories
Data source: 1995 CWSS. Connections include residential, commercial, and industrial connections.

These eliminations resulted in a data set consisting of 1,734 records with paired responses for population and total average daily flow, weighted for item-level non-response.

4.2.2 Distribution Analysis of Regression Variables

The distributions of the average daily flow and population data were evaluated using stem-and-leaf plots, box plots, normal probability plots, coefficients of variation, and the Shapiro-Wilk test statistic. This evaluation showed that it would not be reasonable to assume the data were normally distributed (e.g., plots did not appear normal in shape and the hypothesis of normality was rejected using the Shapiro-Wilk

[7] Only 65 systems in the United States serve a total population greater than 500,000. Because they are few, but have a potentially large impact (large systems serve about 20 percent of the population), EPA plans to perform regulatory analyses for these systems individually.

statistic at the 5 percent significance level). When the data were transformed to a natural logarithm scale, however, the same evaluation showed that it would be acceptable to assume flow and population were log-normally distributed (e.g., plots displayed normality and the hypothesis of log-normality was not rejected at the 5 percent significance level). Based on this result, both the population and flow variables were transformed to a natural logarithm scale (i.e., a log-transformed regression model was used).

4.2.3 Average Daily Flow and Population Regression

Exhibit 4.2 presents the regression for average daily flow versus retail population. The data show a very good correlation, as indicated by the r value (0.90). While the fit is excellent, a limited number of systems lie a considerable distance from the regression. A more detailed analysis of these points was undertaken to determine whether the systems were outliers representing extreme isolated cases, or should be included in developing nationally representative models.

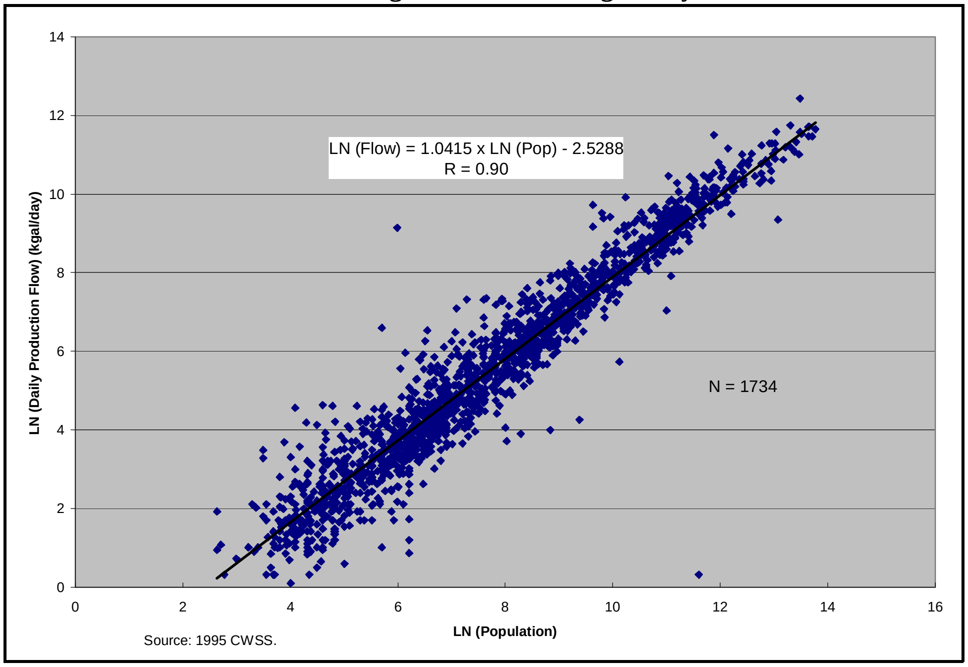
Two approaches were considered for incorporating the variability of each system while eliminating those that represented extreme cases. First, capping flow at the minimum and maximum per capita per day values across the United States from U.S. Geological Survey (USGS) data,[8] which would have resulted in eliminating any system with flows less than 109 gallons per capita per day (gpcd) or greater than 344 gpcd, was considered.[9] Comparison of flows reported in the CWSS with these values indicated that trimming the data set based on the USGS information would have resulted in eliminating about 20 percent of the data. This would have dramatically reduced the sample size. Furthermore, the available literature and engineering judgment indicated that a variability in flow rates much wider than that suggested by the USGS data would be reasonable (Bauman et. al, 1998). That is, only a small fraction of systems have flows that would be considered truly extreme.

The second approach consisted of evaluating data points outside the 95 percent confidence band to determine if they represent typical CWSs. Based on the literature and engineering judgment, systems in this extreme range would encompass only facilities with very “low” or “high” flows. Also, these data points, as observed in Exhibit 4.2, lie a significant distance from the vast majority of data points in the regression. This approach was selected to maintain the representativeness of the data set while reducing the effect on regression analysis from extreme values. 16 of 1,734 systems that reported average daily flow outside the 95 percent confidence band were evaluated for representativeness.

[8] U.S. Geological Survey. 1993. *Estimated Use of Water in the United States in 1990*. Circular 1081.

[9] Note that these USGS values reflect “public supply” and include commercial and industrial uses. Because some small CWSs serve domestic uses only, a more reasonable lower bound might be that for domestic use only. The lowest per capita domestic use rate reported by USGS was 23 gpcd. Even using this lower bound,

Exhibit 4.2. Regression for Average Daily Flow

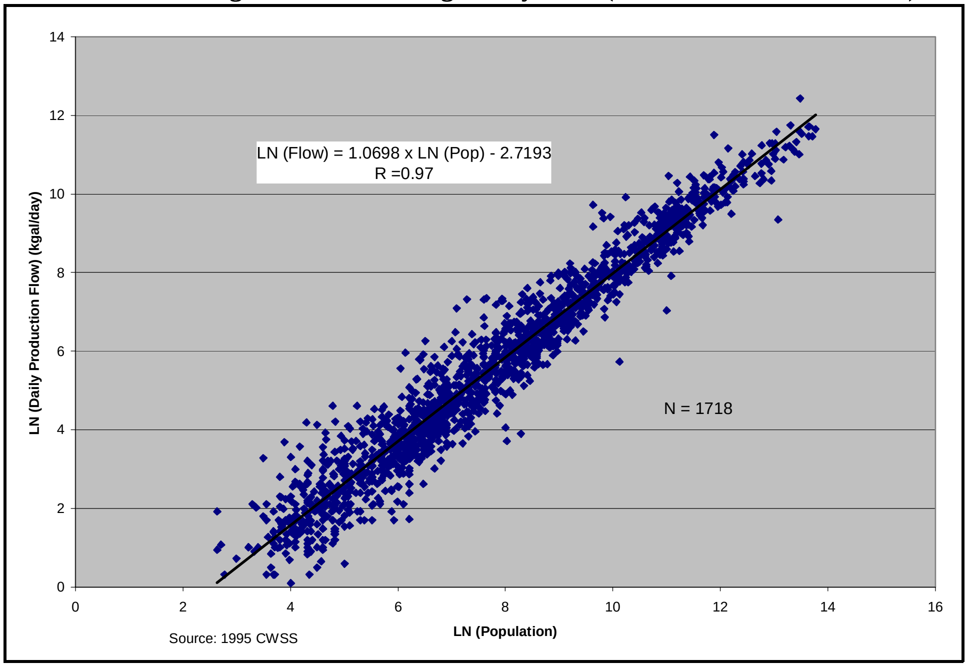


Of these 16 systems, 10 had very low flows, ranging from 0 to about 9 gpcd, and 6 systems had very high flows, ranging from about 1,000 to 23,000 gpcd. Using their specific survey responses, it was possible to draw conclusions about whether some of these 16 systems actually represented CWSs. Three of the low-flow systems wholesaled all or nearly all of their flow and reported a population that appeared to represent client systems, not retail, customers served. Two of the other low-flow systems were mobile home parks and may have represented seasonal use. Other low flow systems may include populations using wells. Of the high-flow systems, one was an abbey, one appeared to be a movie studio, one was an irrigation system, one used most of its flow for agriculture, and one used most of its flow for commercial and industrial uses. Therefore, based on the available information, the very high-flow and very low-flow systems that were removed appear not to represent typical CWSs and were eliminated from the data set.

Exhibit 4.3 depicts the resulting regression after the removal of the extreme values. The r value increases from 0.901 to 0.971 after the implementation rule. Both r values demonstrate excellent correlation but the latter number is probably more representative of the universe of CWSs without biases from systems producing extremely high or low volumes of water.

Note that, after data screening and elimination of extreme values, 1,718 systems remained with valid data for average daily flow and population. To improve the accuracy of subsequent analyses of this smaller data set, adjustments were made to the CWSS sample weights, as described below.

Exhibit 4.3. Regression for Average Daily Flow (Extreme Values Removed)



4.2.4 Adjustment of Sampling Weights for Item-Level Nonresponse

The phenomenon of a system not responding to a question (e.g., not reporting a flow) is known as item-level nonresponse (to distinguish it from primary sample unit nonresponse, which means a system did not respond to the survey at all). When some systems in a stratum do not respond (or their responses are excluded), the proportion of the stratum represented by each respondent changes. In these instances, the accuracy of the analysis can be improved by adjusting samples weights.

After the initial regression analysis presented above, the adjustment for item-level nonresponse was performed for the remaining 1,718 records. The adjustment consisted of further modification of the sampling weights provided with the CWSS results, following the same approach used to account for primary sample unit nonresponse (see Section 3.1.2). Subsequent analyses of average daily flow in this report (e.g., in Section 4.4) used the newly weighted data. Although in this case the adjustment resulted in only minor differences in results, the adjustment for item-level nonresponse was made to maintain the validity of the survey design. For details on item-level nonresponse and the technique used to perform regression analysis using weighted data, see Appendix D.

4.3 Regression of Design Flow Versus Population

Question 5b of the CWSS reports the maximum daily treatment capacity for the system, which is based on a variety of engineering, planning, and design considerations. These include peak daily flow, peak hourly flow, fire-fighting requirements, and population-growth estimates. Maximum daily treatment capacity (design flow) determines the total amount of treatment facilities that may be necessary and is used to estimate capital costs of complying with drinking water regulations. Therefore, a regression analysis was performed with design flow as the dependent variable and population as the independent variable.

Note that peak daily flow also was reported in the 1995 CWSS. Peak daily flow also can be used to estimate capital costs. Initial results for peak daily flow were presented to the Technology Design Panel (TDP) at a workshop in November 1997. The TDP included government (local, State, Federal) and industry representatives with expertise in the technical and regulatory aspects of the water treatment business. The TDP recommended that design flow was more appropriate than peak daily flow for regulatory analysis. In addition, systems responding to the CWSS generally reported design flows that were greater than peak daily flows, further indicating that design flow is more representative of the maximum required treatment capacity. Therefore, model systems were based on design flow (peak daily flow was not analyzed further).

4.3.1 Data Screening

The analysis of population and design flow began with the 1,718 records remaining after the analysis of daily production flow. Like daily production flow, the design flow was manipulated to link it directly to the retail population reported in Question 11. The method used to adjust the design flow differed from the method for daily production flow as follows. An initial review of the design to daily production flow ratios showed that some systems had a design to daily production flow ratio less than one (reported delivery exceeds capacity). Since this is not a desirable design or operating condition, it was inferred that some respondents (i.e., systems categorized as primarily ground or surface water users) included purchased water flow in Question 4 (daily production flow), but not in Question 5b (design flow). This provides one explanation for the low design-to-average ratios. Based on this inference, the following adjustment was made to design flow reported in Question 5B:

$$Design\ Flow = \begin{pmatrix} Question\ 5b\ Design\ Flow \\ + \\ Question\ 4c\ Purchased\ Flow \end{pmatrix} - Question\ 29\ Wholesale\ Flow$$

Wholesale flow was subtracted from the total to link the retail population to retail design flow. Following this adjustment, it was found that 239 systems did not have useable design flows (design flows either were not reported by the respondent or were negative following the adjustments). Therefore, paired design flow and population data for the remaining 1,479 systems were used for the regression analysis.

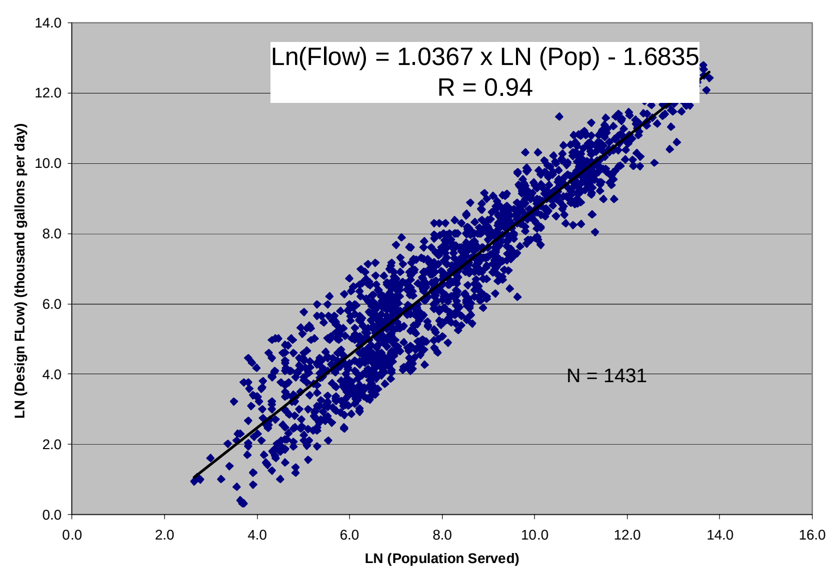
4.3.2 Maximum Daily Treatment Design Capacity and Population Regression

The regression line for design flow and population (Exhibit 4.4) indicates a strong correlation, with an r value of 0.90, comparable to that for daily production flow (Exhibit 4.2). While there are fewer paired data points (1,480) for design flow than for daily production flow, population and design flow continue to show an excellent correlation. There are more systems that reported atypical design flows than daily production flows. As noted previously, this could be the result of design flow being based on variables other than population. An example of a variable affecting design flow is seasonal demand (e.g., for irrigation). Atypical design flows are similar to the extreme flows observed for daily production flows in that they are outside "reasonable" variances used to define baseline characteristics for national-level cost and benefit analysis. To minimize the impact of these atypical or extreme values, systems that reported design flows outside a 95 percent confidence band around the regression line were set aside. As in the previous regression analysis, these extreme values were analyzed separately.

Of the 48 atypical systems, 26 had very low design flows, ranging from less than 1 to about 19 gpcd, and 22 systems had very high design flows, ranging from about 1,500 to 65,000 gpcd. Using the specific survey responses, it was possible to confirm that some of these design flows were atypical of CWSs. For example, the low design flow systems include three mobile home parks, a university, and a country club. Of the high design flow systems, three were mobile home parks and one was an apartment complex. Two more of the systems with high design flows supplied the majority of their water to commercial or industrial customers. Another system supplied all its flow to municipal buildings or parks. Finally, three other systems with high design flows indicated that seasonal demand was more important than or equally as important as current peak needs in determining design flow, suggesting these systems have a high demand for irrigation or have fluctuating seasonal populations (e.g., resort areas). Therefore, based on the available information, these atypical systems appeared not to represent CWSs, and were removed from the analysis. This resulted in a final data set consisting of 1,431 records for subsequent subcategorization (see following sections).

Exhibit 4.4 shows the regression line after the atypical values were removed. Following the elimination of the 48 systems, sample weights were regenerated for the remaining systems using the protocol described in Section 4.2.4 to account for item-level non-responses.

Exhibit 4.4. Regression Line for Design Flow (Extreme Values Removed)



4.4 Regression Analysis of Different Categories of CWSs

The regression analyses of daily production and design flows presented in the previous sections confirm the strong relationship between population and flow. These analyses were based on pooled data that did not distinguish between systems based on water source or ownership.

For regulatory analysis purposes, a model that distinguishes systems by ownership and water source is advantageous in terms of precision and accuracy. For example, a given regulation might apply differently to ground water systems than to surface water systems (e.g., a specific contaminant might be expected to be present only in ground water). The treatment configurations of ground water and surface water systems also vary. Surface water systems have fewer intakes and entry points, while ground water systems can have large networks of wells. Purchased water systems often do not have their own treatment facilities, but buy all treated water from another system. Finally, upcoming regulations address ground and surface water systems separately. Considering these factors, regulatory analyses often must be able to address costs and benefits by water source category.

In addition to source water type, stratifying systems by ownership category may also be advantageous for regulatory analysis purposes. For example, costs for labor and capital can differ for public and private systems. It also was suspected that the additional oversight provided by public utility commissions could affect typical system capacity. For these reasons, it was deemed appropriate to examine systems by ownership category to determine whether subcategorization of flow models was necessary.

Systems were categorized by ownership (public or private) and by source (ground water, surface water, or purchased water), resulting in the following six strata.[10,11]

- Public surface water systems
- Private surface water systems
- Public ground water systems
- Private ground water systems
- Public purchased water systems
- Private purchased water systems

A regression line for each source water type was generated. Then, separate regression lines for each of the six classifications were generated. The lines were tested statistically to determine if the lines are different to provide a statistical basis in addition to regulatory analysis needs.

4.4.1 Regression Analysis for Different Strata

Exhibits 4.5 and 4.6 present the regression equations (weighted for item-level nonresponse) for different CWS categories for daily production flow and design flow, respectively. Based on their r values, all the lines continue to display a strong correlation. This confirms that the strong relationship remains for all of the classifications.

[10] The CWSS also classifies systems as ancillary. These systems produce water as a secondary activity to their primary business function (for example, a paper mill that supplies potable water to its workers or sells it to the public). These systems typically serve small populations (less than 500) and for all practical purposes function like a private utility. Accordingly, they were combined with other private water systems.

[11] As discussed in Section 2.5, approximately 20 percent of surface water systems have some ground water flow. For purposes of this report, these systems were categorized based on the source accounting for the majority of their flow.

Exhibit 4.5. Daily Production Flow Regression Equations for Subcategories of CWS		
CWS Category	**Regression Equation (Daily Production Flow)**	**95% Confidence Interval***
Ground (Public)	$Y = 0.08575 X^{1.05839}$	$\pm 1.965 [0.000651 + (\ln(X)-8.049)^2 /7512.73]^{1/2}$
Ground (Private)	$Y = 0.06670 X^{1.06284}$	$\pm 1.965 [0.000973 + (\ln(X)-6.914)^2 /4746.71]^{1/2}$
Surface (Public)	$Y = 0.14004 X^{0.99703}$	$\pm 1.969 [0.001004 + (\ln(X)-9.334)^2 /4858.16]^{1/2}$
Surface (Private)	$Y = 0.09036 X^{1.03338}$	$\pm 1.976 [0.003628 + (\ln(X)-7.620)^2 /1795.38]^{1/2}$
Purchased (Public)	$Y = 0.04692 X^{1.10189}$	$\pm 1.970 [0.001584 + (\ln(X)-7.954)^2 /3251.88]^{1/2}$
Purchased (Private)	$Y = 0.05004 X^{1.08339}$	$\pm 1.972 [0.001335 + (\ln(X)-6.873)^2 /2388.03]^{1/2}$
* Due to the statistical complexity involved in the calculating weighted confidence intervals, confidence intervals shown are those for the corresponding unweighted regression results. Weighted confidence intervals would be very similar. Notes: Y = daily production flow (thousand gallons per day); X = population served. Regression equations are weighted for item-level nonresponse.		

Exhibit 4.6. Design Flow Regression Equations for Subcategories of CWS		
CWS Category	**Regression Equation (Design Flow)**	**95% Confidence Interval***
Ground (Public)	$Y = 0.54992 X^{0.95538}$	$\pm 1.967 [0.001384 + (\ln(X)-8.335)^2 /3439.07]^{1/2}$
Ground (Private)	$Y = 0.41682 X^{0.96078}$	$\pm 1.968 [0.002070 + (\ln(X)-7.415)^2 /2194.66]^{1/2}$
Surface (Public)	$Y = 0.59028 X^{0.94573}$	$\pm 1.970 [0.001326 + (\ln(X)-9.381)^2 /3540.72]^{1/2}$
Surface (Private)	$Y = 0.35674 X^{0.96188}$	$\pm 1.979 [0.004480 + (\ln(X)-7.948)^2 /1391.46]^{1/2}$
Purchased (Public)	See Section 4.5	
Purchased (Private)	See Section 4.5	
* Due to the statistical complexity involved in the calculating weighted confidence intervals, confidence intervals shown are those for the corresponding unweighted regression results. Weighted confidence intervals would be very similar. Notes: Y = design flow (thousand gallons per day); X = population served. Regression equations are weighted for item-level nonresponse.		

4.4.2 Statistical Tests To Determine Differences in Regression Lines

As noted above, the regression lines were tested to provide a statistical basis for subcategorizing the flow and population model by ownership. First, for each source category, a regression analysis was performed to relate population and daily production flow assuming no differences between systems based on ownership. Then, a “dummy variable” was created to account for ownership category. A dummy variable is an artificial measure created to describe a qualitative factor (in this case, a categorization) that cannot be measured numerically. The dummy variable used here was set equal to 1 for public systems and 0 for private systems. Additional regressions were performed incorporating the dummy variable into the original regression equation in various ways. The effect of ownership category was then assessed by comparing the results of these regressions and the original regression using statistical tests.

For each source category, these statistical tests concluded that regression models for public and private systems are not identical. That is, F-tests comparing the original model to one incorporating the dummy variable in both the slope and intercept term found differences between the two models at the 1 percent significance level for ground and surface water systems and at the 10 percent significance level for purchased water systems. Further tests examined the nature of these statistical differences. While these tests did not demonstrate that ownership has a significant effect on the slope of the regression line, they did conclude that ownership significantly affects the intercept (i.e., causes a parallel shift, up or down, of the flow). That is, F-tests comparing the original model to one incorporating the dummy variable into the intercept term found differences at the 1 percent significance level for ground and surface water systems and at the 5 percent significance level for purchased water systems. These results were corroborated by comparing the least square means of private and public systems within each water source category. This second set of tests showed a statistical difference between ownership categories at the 5 percent significance level for all three source water categories. The statistical calculations are provided in Appendix D.

Once the statistical tests for daily production flow supported subcategorization, the same source and ownership stratification was extended to design flows. Inconsistent subcategories among daily production flow and design flow could result in a incongruous (and inconvenient) subcategorization scheme.

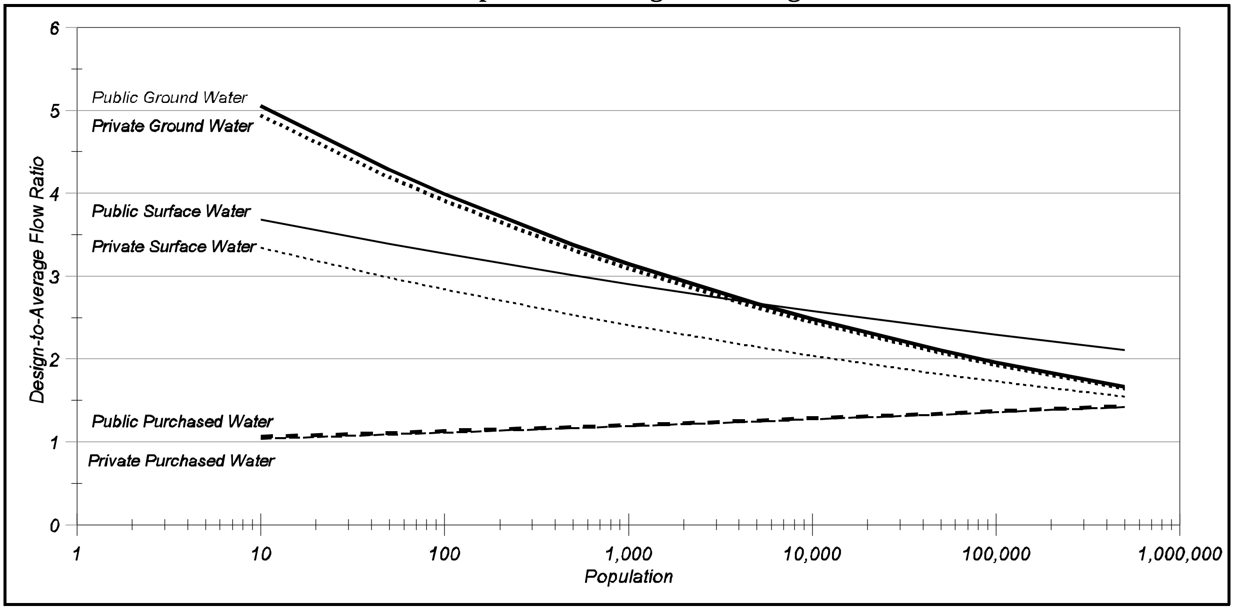
4.5 Evaluations of Design-to-Average Flow Ratios

While the regression results for both design and average (daily production) flows are quite strong, it is also important to consider the relationship between the two flows to confirm the reasonableness of the full design capacity, which is necessary in water systems to ensure adequate service during peak demand periods, for emergency flows, and for seasonal demand. EPA consulted the Denver Technology Design Panel on the subject of full design capacity needs.

The Panel recommended that all water systems should have a minimum of 100 percent design capacity (meaning the ratio of design flow to average flow should be at least 2). Further, it was believed that more design capacity would be required by smaller systems. The Panel, however, did not reach agreement on a specific minimum design capacity for these systems.

Design-to-average flow ratios were evaluated by source water type (ground, surface, purchased) for the range of flows being modeled. In general, design-to-average ratios decrease with increasing system size because of differences in treatment plant and distribution system configuration and water demand. Smaller systems typically use less storage and experience sharper peaks in demand. Accordingly, they are expected to have higher design-to-average ratios (to meet sharper peaking factors) than larger systems. Exhibit 4.7 shows design-to-average ratio plots for the six categories of CWS. Both public and private ground water systems have design-to-average ratios ranging from about 1.5 to 4.3. Surface water systems have lower design-to-average ratios than ground water systems. Also, public surface water systems have higher design-to-average ratios than private surface water systems. This difference may stem from public surface water systems adding extra capacity based on a longer planning horizon, or from private systems having more frequent meter replacement and better control of unaccounted for water.

Exhibit 4.7. Comparison of Design-to-Average Flow Ratios



This analysis of design-to-average flow ratios suggests that some modifications to the design flow and population model derived in the previous section are necessary for regulatory analysis purposes. The reasoning behind and details of these modifications are discussed in the following sections.

4.5.1 Design Flow Modification for Surface and Ground Water Systems

The Denver Technology Design Panel suggested to EPA that, for regulatory analysis purposes, a minimum design-to-average ratio of 2 should be used for large systems. The panel indicated it was unlikely that systems would install new capacity below this ratio. The population where the regression equations produce design-to-average flow ratios less than 2 varies by CWS type. The equations in Exhibit 4.8, therefore, should be used for estimating design flow for the ranges of population shown. The equations in Exhibit 4.8 were developed using the following approach:

(1) For populations for which the design-to-average flow ratio is greater than 2, use the design flow equation resulting from the regression analysis in Section 4.4.

(2) For populations for which the design-to-average flow ratio is less than 2, the daily production flow equation (from the regression analysis in Section 4.4) was modified to produced a design flow twice as large.

Exhibit 4.8. Recommended Approach for Estimating Design Flow for Surface and Ground Water Systems

Category	Recommended Approach to Estimate Design Flow
Public Ground	$X < 100{,}000$. $Y = 0.54992\ X^{0.95538}$ $X > 100{,}000$. $Y = 0.17150\ X^{1.05839}$
Private Ground	$X < 90{,}000$. $Y = 0.41682\ X^{0.96078}$ $X > 90{,}000$. $Y = 0.13340\ X^{1.06284}$
Public Surface	For all populations. $Y = 0.59028\ X^{0.94573}$
Private Surface	$X < 20{,}000$. $Y = 0.35674\ X^{0.96188}$ $X > 20{,}000$. $Y = 0.18072\ X^{1.03338}$

Note: Y = design flow (thousand gallons per day); X = population served

4.5.2 Design Flow Modification for Purchased Water Systems

As described in Section 4.2, systems are categorized by retail flow and population. This is done by removing the wholesale portion of flow and assigning it to purchased water systems. It is assumed that purchased water systems would bear the full cost of the design flow of their parent water system and would exhibit similar design-to-average flow ratios. Preliminary regression analyses of the design flows reported by purchased water systems, however, showed design-to-average flow ratios less than 1.5 for all populations, as shown in Exhibit 4.7. The design flow equations were therefore modified for Exhibit 4.8 using the following approach:

(1) Starting with the population of the purchased water system, calculate the daily production flow for the system using the equations developed in Section 4.4.

(2) To estimate a design flow with a similar design to average flow ratio as the parent water system, back-calculate a “virtual” population for an appropriate parent water system (e.g., a private ground water system) using the flow from step 1 above and the average daily flow equation for the parent water system. That is, for a system purchasing water from a private ground water system, substitute the purchased average daily flow into the private ground water daily flow equation and solve for a corresponding “virtual” population.

(3) Use this “virtual” population, calculate a theoretical design flow for the purchased water system using the equations presented in Exhibit 4.6. That is, for a system purchasing water from a private ground water system, use the “virtual” population derived in Step 2 in the equation in Exhibit 4.6 to estimate design flow.

Exhibit 4.9. Estimated Design Flow for Purchased Water Systems	
Corresponding Source Water Category	**Theoretical Estimated Design Flow for Purchased Water System**
Public Ground	$X < 109{,}000$. $Y = 0.3191\ X^{0.9946}$ $X > 109{,}000$. $Y = 0.09384\ X^{1.10189}$
Private Ground	$X < 94{,}000$. $Y = 0.3215\ X^{0.9794}$ $X > 94{,}000$. $Y = 0.10008\ X^{1.08339}$
Public Surface	For all populations. $Y = 0.2092\ X^{1.0452}$
Private Surface	$X < 21{,}000$. $Y = 0.2058\ X^{1.0084}$ $X > 21{,}000$. $Y = 0.10008\ X^{1.08339}$
Note: Y = theoretical design flow (thousand gallons per day); X = purchased water system population served	

4.6 Summary of Population and Flow Relationships

Exhibits 4.5 through 4.9 detail the equations derived in this analysis for average daily flow and design flow for the various categories of CWSs. Exhibits 4.10 through 4.13 present the results of this analysis graphically. Using the model developed here, as shown in these figures, derived design flows are always greater (at least 2 times) than derived average daily flows. On the logarithmic scale presented here, flows for private systems are generally (but not always) lower than those for public systems. In the Exhibits, differences between private and public systems may appear small, but these differences are statistically significant, as discussed in Section 4.4.2. For example, a ground water system serving 100 people would be projected to have a design flow of 45,000 gallons per day if it were a public system versus a design flow of 35,000 gallons per day if it were a private system.

Exhibit 4.10 Population and Flow Relationships for Ground Water

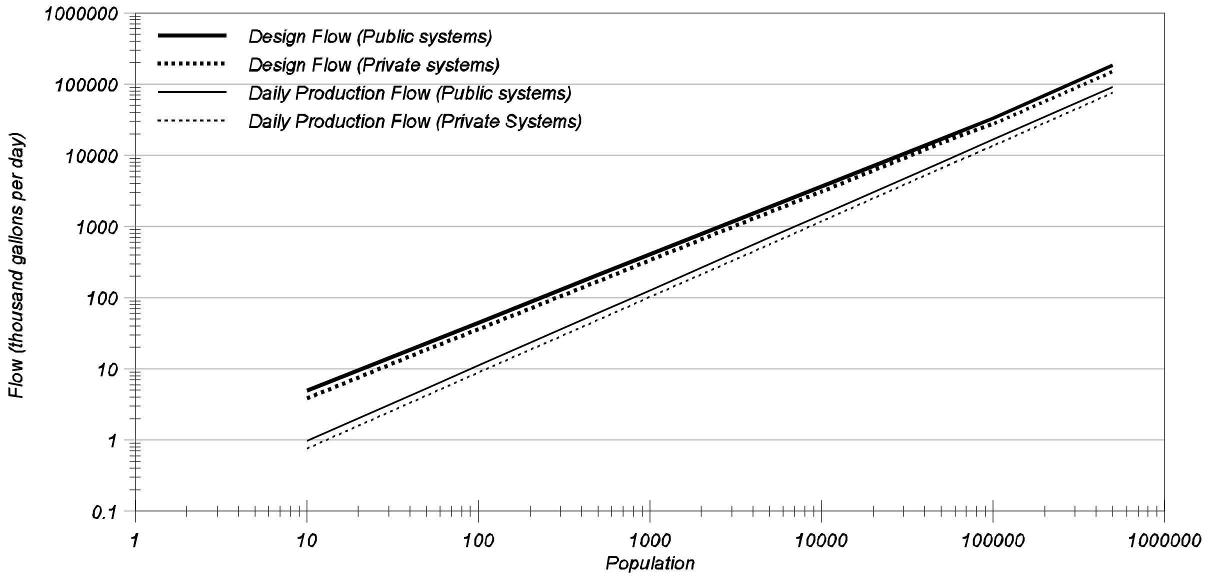


Exhibit 4.11 Population and Flow Relationships for Surface Water

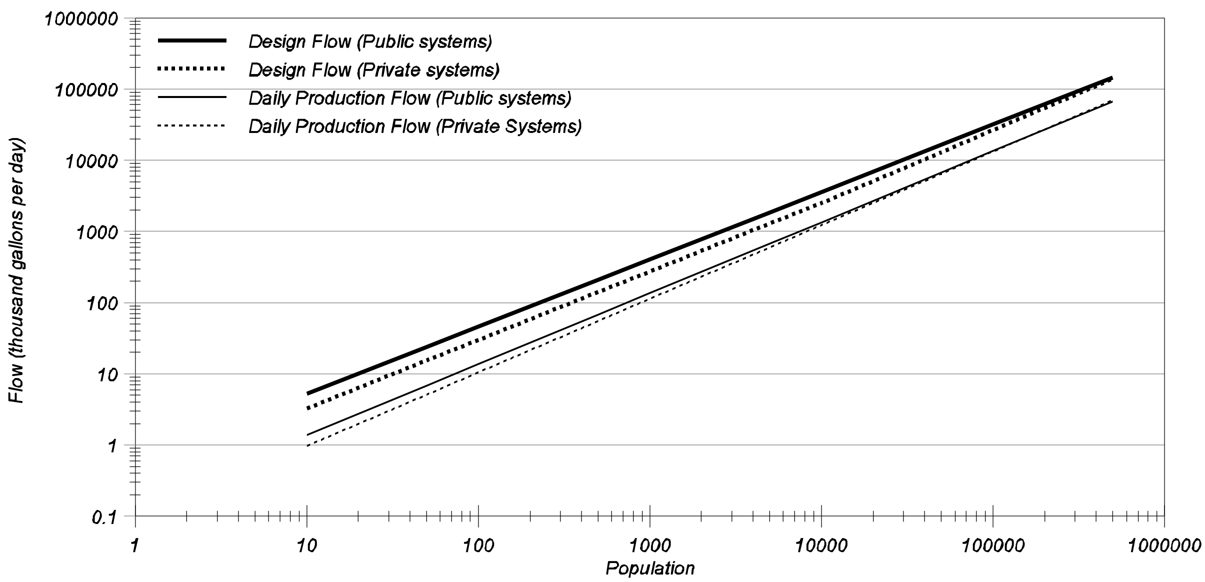


Exhibit 4.12 Population and Flow Relationships of Purchased Water Systems Fed by Ground Water

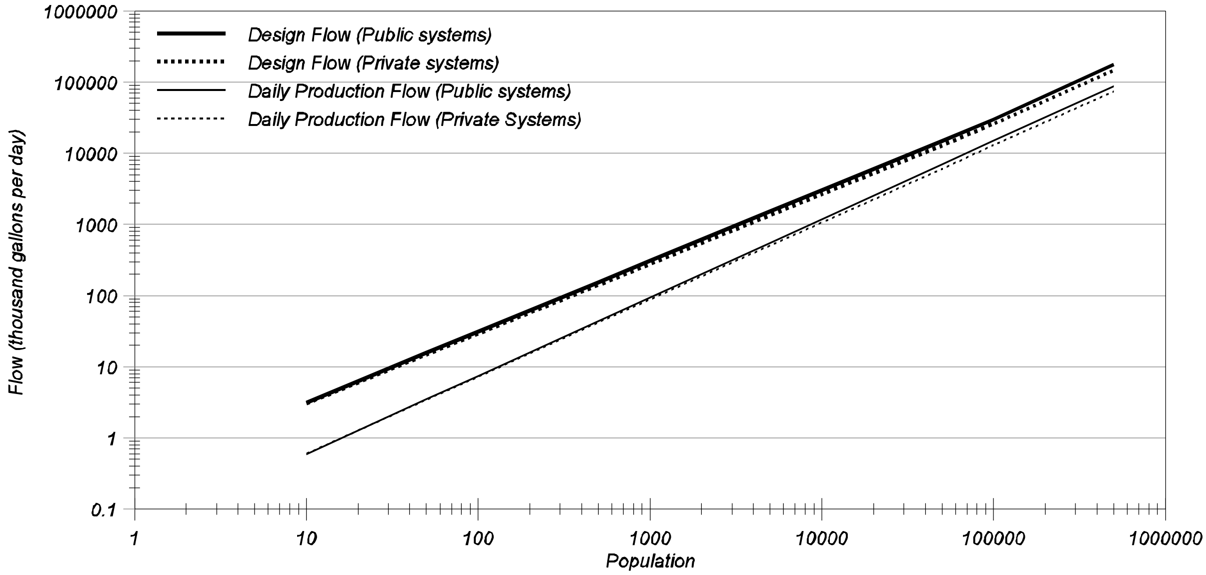
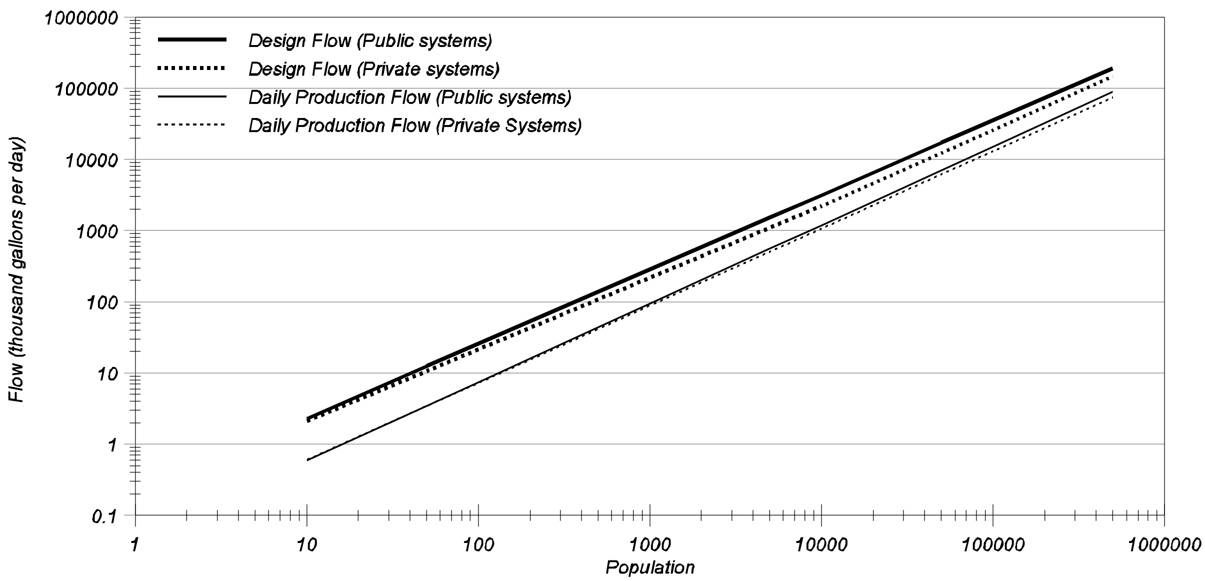


Exhibit 4.13 Population and Flow Relationships for Purchased Water Systems Fed by Ground Water



5: Analysis of Entry Point Configurations

Entry points are locations where untreated water, treated water, or purchased water enter the distribution system network. A public water system may just have one entry point supplying all of its drinking water, or multiple entry points from different types of sources. In a system with more than one entry point, one may provide a majority of the flow to the distribution system. Generally, larger systems have more complex configurations. Also, configurations can be more complex for ground water systems because individual treatment plants can be supplied by networks of wells. Ground water systems also can have untreated wells (or networks of untreated wells) connected to the distribution system. Exhibit 5.1 presents examples of surface and groundwater entry point configurations.

In Chapter 4, a design and average flow model was developed at the system level. This chapter expands on that model by providing the information necessary to address systems based on the number of entry points. For regulatory impact analysis purposes, it may be relevant to distinguish systems with multiple entry points. For example, consider a ground water system with multiple wells. If a regulated contaminant affects only one ground water well, then treatment would be required only for the entry point using that well. An accurate estimate of compliance costs for the system would consider treatment only at that entry point. Also, the regulatory impact analysis may need to consider differences in economies of scale between systems that treat all water in a common facility versus separate facilities at separate entry points.

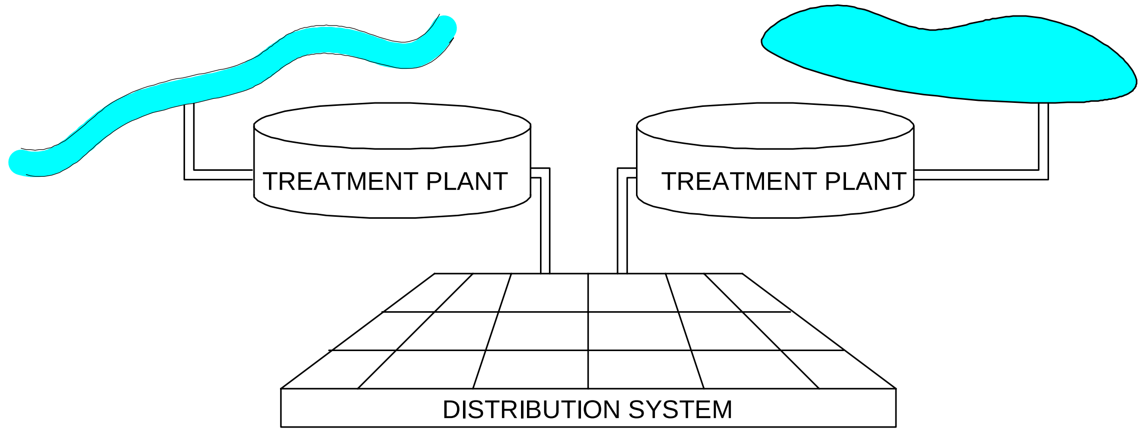
For systems with multiple entry points, other aspects of their configurations beyond the number of entry points also may be relevant for regulatory analysis purposes. Because estimates of compliance cost typically are based on flow rates, the distribution of flow across entry points may be relevant. For ground water systems, impacts may be limited to individual wells, depending on their depth. For example, immobile contaminants may affect shallow wells only.

The purpose of this chapter is to analyze the characteristics of entry points for ground and surface water systems. Specifically, this chapter analyzes the following:

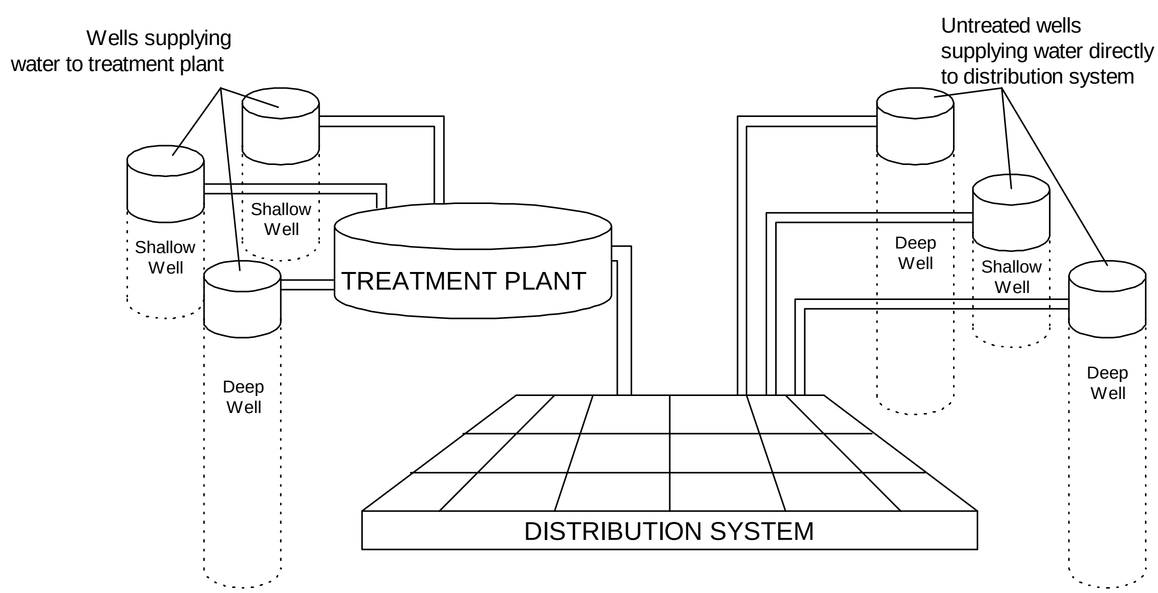
- The numbers of entry points for ground and surface water systems [12]
- The distribution of flows among entry points
- The spatial distribution of entry points
- The depths of ground water wells.

[12] Ownership distinctions (i.e., public and private) are not made for ground and surface water systems because disaggregation at this level would severely limit the number of data points available for analysis of entry points.

Exhibit 5.1. Conceptual Diagram of Entry Point Configurations



A surface water PWS with two treatment plants



A ground water PWS with a treatment plant and untreated wells

5.1 Data Cleaning

Data for entry points are reported in Questions 18 and 20 of the CWSS (Appendix C). Each row in Question 18 represents a treatment facility. Question 18 of the CWSS questionnaire reports latitudes and longitudes for each treatment plant, number of wells treated (if it is a ground water plant), range of depth, flows, and treatments provided. Question 20 provides information for each well or surface water intake not receiving treatment. Question 20 reports the aquifer or surface water source name and type (ground or surface), location of each well and its depth, and flow.

For ground water systems, all of the wells reported in Question 18 are connected to treatment plants, while the wells reported in Question 20 are connected directly to the system's distribution network. For surface water systems, Question 18 reports intakes connected through treatment plants, while Question 20 reports untreated intakes connected directly to the distribution network. Thus, the sum of the rows in Questions 18 and 20 corresponds to the number of entry points for each system.

Questions 18 and 20, which were not validation fields in the CWSS, were completed less frequently than the validated fields (e.g., Question 4). Any system that did not report well and/or flow information for an entry point in Questions 18 and 20 was eliminated. Purchased water systems also were eliminated from the data set, since their configurations are not representative of those of stand-alone systems. The data cleaning process resulted in two data sets comprising 840 ground water systems with 2,249 entry points based on Questions 18 and 20 and 376 surface water systems with 476 entry points based on Questions 18 and 20. These data sets were used to analyze entry point characteristics.

5.2 Number of Entry Points

For each of the eight CWSS population size categories described in Chapter 3, Exhibit 5.1 shows the frequency with which the sampled ground and surface water systems have multiple entry points. The data in Exhibit 5.1 generally follow the expected trend. That is, the smaller systems tend to have a single entry point and larger systems tend to have multiple entry points. Often, smaller systems can meet the demand with a single source.

For ground water systems, Exhibit 5.1 indicates that the majority of systems sampled in the two smallest population categories have a single entry point. About one-third of the systems in the 501 to 1,000 category, which are considered small systems, have more than one entry point. In the next two population categories, the percentages of systems with more than one entry point are 42 percent and 54 percent, respectively. The use of multiple entry points increases as systems get larger. The reliance on multiple entry points by a significant number of small systems is relevant for regulatory impact analyses because of issues discussed above (e.g., situations in which contaminant impacts are entry point specific, rather than system wide).

The pattern for surface water systems is different. Multiple entry points become an issue for systems serving more than 3,300 persons. Furthermore, with respect to the number of entry points, surface systems did not report as great a variety of configurations as ground water systems. Even for the large population categories, the majority of surface water systems reported one or two entry points, with a maximum of six reported. By comparison, groundwater systems reported values ranging from 1 to more than 30 entry points. To further characterize the number of entry points for systems of various population sizes, EPA performed additional statistical analyses as presented below.

Exhibit 5.2. Frequency of Multiple Entry Points						
	Ground Water Systems			Surface Water Systems		
Population Category	Percent with One Entry Point	Percent with Multiple Entry Points	Maximum Number of Entry Points	Percent with One Entry Point	Percent with Multiple Entry Points	Maximum Number of Entry Points
Less than 100	86.9%	13.1%	4	100.0%	0.0%	1
101 to 500	80.5%	19.5%	11	97.7%	2.3%	2
501 to 1,000	66.9%	33.1%	4	100.0%	0.0%	1
1,001 to 3,300	58.1%	41.9%	13	98.1%	1.9%	2
3,301 to 10,000	46.4%	53.6%	23	89.8%	10.2%	4
10,001 to 50,000	33.0%	67.0%	18	91.2%	8.8%	2
50,001 to 100,000	26.5%	73.5%	37	59.2%	40.8%	6
100,001 to 1,000,000	24.2%	75.8%	31	45.2%	54.8%	4

Ground water systems even small systems, reported a wide range of configurations with respect to number of entry points. Mean and percentile values for the number of entry points were calculated for ground water systems in each population category. Because the number of systems samples in each population category was relatively small (compared to those in previous chapters0, these statistics were generated using a computer-intensive statistical procedure[13]. Using the bootstrap estimates, Exhibit 5.2 characterizes the mean number of entry-points and the percentile distribution of the number of entry points for ground water systems in each population category. The percentile data in Exhibit 5.2 are the "typical" (i.e., bootstrap mean) number of entry points for systems in the xth percentile. For example, these data may be interpreted as follows: The typical number of entry points for systems in the 75th percentile of all systems serving 100 to 500 people is one. Appendix F presents a more detailed breakdown of the bootstrapping results.

Because surface water systems did not report as great variation in the number of entry points, similar detailed characterization was not necessary. As discussed above, even for the large population categories, the majority of surface water systems reported one or two entry points. Recently collected data from Information Collection Request for large surface water systems supports this estimate.

[13] This procedure, known as "Bootstrapping", allows statistical estimates to be generated from smaller sample sizes with nonnormal distributions without the need for extensive assumptions. The bootstrap method draws a large number of random samples (in this case 10,000) with replacements and calculates the statistics of interest for each sample. Item nonresponse factors and adjusted weights, as discussed in previous chapters, were used in the bootstrap analysis.

Exhibit 5.3. Percentile Distribution of Number of Entry Points for Ground Water Systems						
Population Category	**Number of Entry Points***					
	Mean	**10th Percentile**	**25th Percentile**	**50th Percentile**	**75th Percentile**	**90th Percentile**
Less than 100	1	1	1	1	1	2
101 to 500	1	1	1	1	1	2
501 to 1,000	2	1	1	1	2	3
1,001 to 3,300	2	1	1	1	2	4
3,301 to 10,000	2	1	1	2	3	4
10,001 to 50,000	4	1	1	3	5	8
50,001 to 100,000	6	1	1	4	8	17
100,000 to 1,000,000	9	1	1	5	15	24
* Bootstrap value, rounded to the nearest integer.						

5.3 Distribution of Flow Among Entry Points

For systems with multiple entry points, this section evaluates the distribution of the total system flow among the entry points. The purpose of this evaluation was to determine if flows are distributed evenly for systems with multiple entry points or if one of the entry points accounts for a majority of the system's total flow.

The distribution of flow among entry points can be significant for regulatory analysis purposes. For example, in a case where a single entry point is affected for a system having three entry points, compliance costs would be estimated based on flow for the affected entry point. The costs would be greater if the affected entry point accounted for a majority of the flow, as opposed to an equal share (one-third) of the flow.

Systems reporting multiple entry points were examined further for the distribution of flow across entry points. The ratios of entry point flow may differ during peak production, however the CWSS does not provide data on peak flow from individual water sources. Therefore for each system, the percentage of total flow accounted for by each entry point was calculated by comparing the entry point's average daily flow to the system's total average daily flow. Entry points were ordered according to their percent contribution to flow (i.e., first entry point = the largest, etc.). Systems were grouped by water source and number of entry points.[14] For each group, arithmetic mean percentages were estimated for each entry point in the ordered

[14] Further subcategorization of systems by population category resulted in too few samples in each group to generate statistics with a high degree of confidence. Examination of the available data did not support the conclusion that there are significant differences between population categories in distribution of flow across entry points. Thus, systems were grouped across all population categories.

set of entry points (i.e., mean percent for the largest entry points, mean percent for the second largest entry points, etc.). Exhibits 5.3 and 5.4 report the mean distribution of flow across entry points for ground and surface water systems, respectively.

Exhibit 5.4. Distribution of Flow by Entry Point (Ground Water Systems)

Number of Entry Points (EP)	Number of Sample Systems	Percent of Total Flow at...					
		Largest EP	2nd Largest EP	3rd Largest EP	4th Largest EP	5th Largest EP	6th Largest EP
2	136	64.7	35.3	-	-	-	-
3	67	50.5	30.4	19.1	-	-	-
4	42	40.5	26.7	19.0	13.8	-	-
5	25	36.7	24.4	16.0	13.0	9.9	-
6	12	42.3	18.1	15.4	10.1	8.3	5.9

Data are shown only for systems having up to 5 entry points. Similar data was generated for systems with up to 37 entry points and is presented in Appendix G.

Exhibit 5.5. Distribution of Flow by Entry Point (Surface Water Systems)

Number of Entry Points (EP)	Number of Sample Systems	Percent of Total Flow at...				
		Largest EP	2nd Largest EP	3rd Largest EP	4th Largest EP	5th Largest EP
2	40	63.1	37.0	-	-	-
3	9	48.6	31.1	20.2	-	-
4	7	45.3	27.1	16.4	11.2	-
6	1	43.4	24.2	16.2	6.1	6.1/4.0*

*4 percent for the sixth largest entry point
Note: None of the sample surface water systems reported exactly five entry points

Data for ground water systems (Exhibit 5.3) show that flows are not evenly distributed across entry points. For systems with two entry points, flows are concentrated at one entry point. When there are more than two entry points, flows appear to be concentrated at a few entry points. As the number of entry points increases, the relative contribution of subsequent wells becomes increasingly smaller.[15] That is, the distribution is skewed. This suggests that it may be desireable for regulatory analysis tools to

[15] Exhibit 5.3 shows data for ground water systems with two to five entry points. Data are available for systems with more than five entry points; however, after five entry points the number of samples drops to 12 systems reporting six entry points. For more than 10 entry points, the number of samples is in the single digits for each category.

consider whether treatment is needed for all entry points. The practical significance of these data will be evaluated as part of sensitivity analysis.

Exhibit 5.4 shows similar data for surface water systems. Flow for surface water systems is distributed in almost the same proportions as ground water systems for a comparable number of entry points. For example, for two entry points, the split for surface and ground water systems are about the same for the largest entry points. For six entry points, the split is comparable. Unlike ground water systems, there are fewer data points (sample systems) and the maximum reported number of entry points is six (compared with up to 37 for ground water systems).

5.4 Spatial Distribution of Entry Points

As discussed in the introduction, when contamination is localized, only a single entry point in a system may be affected. Which entry points are affected by the presence of contamination in the environment depends on their proximity to the contaminated area. In some cases, the number of entry points affected, and therefore compliance costs, may depend on the distance between entry points. In addition, systems facing localized contamination may have the option of shutting down affected sources and drawing water from unaffected (hydrologically separate) entry points. Provided the distant entry points have sufficient capacity and transmission costs are reasonable, this practice can serve as a less costly alternative to installing treatment.

Questions 18 and 20 provide latitude and longitude data for the location of each entry point. For systems with two entry points, a preliminary analysis was conducted by entering these data into a Geographic Information System (GIS), which was then used to calculate the distance between entry points. A limited number of data points were available for this analysis. Twenty-six surface water and 56 ground water systems with two entry points provided latitude and longitude data.[16]

Calculated distances between entry points ranged from 0 meters to 67 kilometers for ground water systems and from about 2 kilometers to 106 kilometers for surface water systems. The upper bounds of both of these ranges appear anomalous and may be the result of inaccurate latitude and longitude data. The majority of distances were between a few hundred meters and 9 kilometers for ground water systems and between 2 kilometers and 30 kilometers for surface water systems. There were too few data points to examine differences between population categories.

The GIS data were also used to analyze the nine surface water systems with more than two entry points that provided latitude and longitude data. For each of these systems, the entry point coordinates were used to define a polygon and the distance from each entry point to the polygon's centroid was calculated. This analysis yielded a range of calculated distances similar to that above, but also revealed a number of anomalous data points. (For example, based on the entry point coordinates, one system appeared to span an area nearly the size of Pennsylvania.) Due to data and other limitations about the accuracy of the latitude and longitude data, a similar analysis was not conducted for the 68 ground water systems with more than two entry points that provided latitude and longitude data.

---

[16] Latitude and longitude coordinates for a few entry points appeared to be transposed; these data points were corrected prior to analysis. Entry points with coordinates outside the United States were deleted.

Thus, preliminary analyses suggest a wide variation in spatial configurations for entry points. They confirm that distant entry points can exist for both ground and surface water systems; however, the available data are too limited to accurately quantify differences between systems. Furthermore, regulatory impacts are likely to be highly system-specific, depending not only on distance between entry points, but also on the nature of the contamination and available supply alternatives.

5.5 Well Depths

Referring back to Exhibit 5.1, ground water systems with multiple wells can draw water from various depths. In some cases, an immobile ground water contaminant may affect only shallow wells. Thus, the number of entry points affected, and, therefore, compliance costs, may depend on well depth. In addition, when immobile contaminants are present, systems may have the option of shutting down shallow wells and drawing water from deeper, unaffected wells. Provided the deep wells have sufficient capacity, this practice can serve as a less costly alternative to installing treatment.

Questions 18 and 20 provide well depth information. In Question 20, depths are reported for individual untreated wells. In Question 18, the system may report multiple wells connected to each treatment facility. Where multiple wells are connected to a treatment facility, the depth data are in the form of a range of depths for the wells connected, as opposed to a depth for each individual well. Three different approaches were used to convert the Question 18 ranges to point estimates so that these data could be examined along with the Question 20 individual depths. The three approaches, respectively, used the minimum of each range, the maximum of each range, and the midpoint of each range as point estimates of depth. When the Question 18 and Question 20 data were combined using these approaches, a relatively large number of data points were available for analysis (2,249 entry points for 840 ground water systems).

Differences in well depths among population categories were examined by calculating the mean depth of all entry points in each category using each of the three approaches to estimating depth. Exhibit 5.5 shows the results of this analysis. Based on the category means, entry points in the smallest population category appear to have much shallower minimum, midpoint, and maximum depths. Systems in the largest population category appear to have deeper well depths. Larger systems also may have greater ability to switch to deeper wells in response to contamination.

The presence of very deep wells in some systems, however, suggests that, faced with immobile contaminants affecting shallow wells, these systems may be able to switch to deep, uncontaminated wells. Examining the range of variation of well depths reported by individual systems (i.e., comparing the minimum well depth reported by a system to the maximum well depth reported by the same system) supports this hypothesis. Exhibit 5.6 shows, by population category, the frequency with which systems reported various ranges of well depths. This analysis shows that the majority of systems in the smaller population categories do not have large depth ranges. That is, wells in most of these systems tend to be at similar depths. Therefore, these systems are not likely (based on depth alone) to have the option of switching wells to avoid contaminants.

Exhibit 5.6. Mean Entry Point Depths by Population Category					
			Mean of Entry Point... (depth in feet)		
Population Category	Number of Entry Points	Number of Systems	Minimum Depth	Midpoint Depth	Maximum Depth
Less than 100	116	99	111	120	128
101 to 500	220	159	272	300	328
501 to 1,000	170	115	254	281	307
1,001 to 3,300	256	136	378	408	438
3,301 to 10,000	335	140	386	410	435
10,001 to 50,000	452	109	351	382	414
50,001 to 100,000	387	49	275	309	344
100,000 to 1,000,000	313	33	459	589	718

Exhibit 5.7. Frequency of Variation of Well Depths by Population Category						
	Percent of Systems with Variation of Well Depths (in feet)...					
Population Category	Less than 50	50 to 100	100 to 200	200 to 500	500 to 1,000	Greater than 1,000
Less than 100	75.0%	15.0%	10.0%	0.0%	0.0%	0.0%
101 to 500	64.6%	12.5%	13.2%	6.9%	2.8%	0.0%
501 to 1,000	65.4%	8.7%	10.6%	10.6%	5.8%	0.0%
1,001 to 3,300	50.0%	9.5%	15.1%	17.5%	5.6%	2.4%
3,301 to 10,000	37.8%	14.2%	20.5%	18.9%	5.5%	3.1%
10,001 to 50,000	28.2%	9.7%	14.6%	27.2%	14.6%	5.8%
50,001 to 100,000	17.0%	8.5%	14.9%	27.7%	21.3%	10.6%
100,000 to 1,000,000	16.1%	12.9%	12.9%	29.0%	9.7%	19.4%

6: Analysis of Treatment-in-Place

This chapter presents the approach and results of an analysis of treatment-in-place (existing treatment technologies) at existing CWSs. Section 6.1 presents the rationale for the analysis. Section 6.2 provides a discussion of the data used for the analysis, including the data cleaning performed. A summary of the results of the analysis is provided in Section 6.3.

6.1 Rationale for Treatment-in-Place Analysis

To estimate potential costs incurred by public water systems as a result of future revisions to drinking water regulations, EPA needs to know the types of treatment currently in place at existing water systems. Having that information allows EPA to more accurately estimate costs associated with treatment plant modifications and upgrades that would be necessary for compliance.

To model treatment-in-place, EPA used responses to two questions in the CWSS, namely, Questions 18 and 20. Question 18 requested basic information for the treatment facilities within each CWS. Specifically, responses to Question 18 identify the source of raw water treated (i.e., ground or surface), daily maximum and average flows, and the type of treatment provided at each plant within the CWS. Question 20 identifies *untreated* wells and surface water intakes, along with flows from these entry points.

As presented in the CWSS questionnaire, Question 18 requested that respondents identify treatment technologies using specific water treatment codes (as presented here in Exhibit 6.1). Many existing technologies, such as ozone and chlorine dioxide were in limited use at the time of the survey. Consequently, there would be a high degree of uncertainty in estimates based on only a handful of responses. To simplify the treatment-in-place model and reduce uncertainties in the data, subcategories of treatment technologies that address common treatment issues (i.e., water treatment codes) were grouped into classes considered significant for future EPA rulemakings. For example, a facility using chlorine and another using chlorine dioxide are both identified as performing “disinfection.” Similarly, a facility that performs coagulation and flocculation using aluminum salts is grouped together with a facility that uses polymers in the coagulation and flocculation process. The combined headings of these groupings are presented in Exhibit 6.2.

A number of treatment plants identified a treatment technology as “other.” For these facilities, EPA reviewed CWSS questionnaire submissions to determine whether these data represented any major treatment category that should be characterized in the model system evaluation. This analysis did not identify any additional treatment categories that should be included in the analysis.

Exhibit 6.1. CWSS Water Treatment Codes (Question 18)

Code	Treatment	Code	Treatment
01	Raw water storage	27	Reverse osmosis
02	Presedimentation	28	Pressure filtration
03	Aeration	29	Other filtration
		30	Filtration combinations
	PRE-DISINFECTION/OXIDATION:		
04	Chlorine		ORGANICS REMOVAL:
05	Chlorine dioxide	31	GAC adsorption post contactors
06	Chloramines	32	GAC adsorption filter adsorbers
07	Ozone	33	PAC addition
08	Potassium Permanganate	34	Ion exchange
09	Pre-disinfection/oxidation combinations	35	Air stripping
10	Lime/Soda ash softening	36	Organics removal combinations
11	Recarbonation with carbon dioxide		
			POST-DISINFECTION:
	IRON AND MANGANESE REMOVAL:	37	Chlorine
12	Green sand filtration	38	Chlorine dioxide
13	Chemical oxidation filtration	39	Chloramines
14	Aeration filtration	40	Post-disinfection combinations
		41	Fluoridation
	FLOCCULATION/COAGULATION	42	Hypochlorination
15	Aluminum salts		
16	Iron salts		CORROSION CONTROL:
17	pH adjustment	43	pH adjustment
18	Activated silica	44	Alkalinity adjustment
19	Clays	45	Corrosion inhibitors
20	Polymers	46	Corrosion control
21	Other flocculation/coagulation		
22	Flocculation/coagulation combinations		OTHER TREATMENTS NOT ELSEWHERE CLASSIFED:
	FILTRATION:	47	Other treatment
23	Slow sand		
24	Rapid sand		
25	Dual/Multi-media		
26	Diatomaceous earth		

Exhibit 6.2. Treatment Code Groups

Disinfection:	04, 05, 06, 07, 08, 09, 37, 38, 39, 40, 42	PAC:	33
Aeration:	03, 14, 35	Filtration:	23, 24, 25, 26, 28, 29, 30
Oxidation (Fe/Mn):	12, 13	Coagulation/Flocculation:	15, 16, 19, 20, 21, 22
Ion Exchange:	34	Lime/Soda Ash Softening:	10
Reverse Osmosis:	27	Recarbonation:	11
GAC:	31, 32		

The numbers refer to the treatment category numbers in the CWSS as presented in Exhibit 6.1.

6.2 Identification of Core Data Set

This analysis used the same data set as the entry point analysis in Chapter 5 (i.e., purchased water systems and nonrespondents to Questions 18 and 20 were not included). To address facilities that failed to adequately respond to Questions 18 and 20, an item-level nonresponse weighting factor, as discussed in Chapter 4, was incorporated into the analysis. The incorporation of item-level nonresponse weighting allows the results for the systems that adequately responded to be representative of the full survey population. This approach assumes that non-respondents are not statistically different from respondents in terms of treatments implemented. Review of responses to core validated questions, financial, and other elements of the questionnaire suggest this is reasonable.[17]

Additionally, a facility with multiple "hits" in one category would count as one treatment technology only. For example, a facility reporting water treatment codes of 23 (slow sand filtration), 24 (rapid sand filtration), and 29 (other filtration) counts as only one facility using filtration in the tabulation of treatments-in-place.

6.3 Summary of Results

To be consistent with the models of population and flow, the treatment-in-place analysis results are presented by source water type. The results are also presented by population size category. As discussed above, the analysis used sample weights adjusted for item-level non-response.

Exhibits 6.3 and 6.4 (for ground water and surface water systems, respectively) present the frequency, by population category, with which systems have treatments in each of the categories defined for this analysis. As such, any one treatment facility within a system that reported a specific treatment classifies the system as having that treatment technology. Systems with two treatment plants at one system or systems with two treatment technologies within the same treatment plant (e.g., pre- and post-disinfection) are counted only once.

Exhibit 6.5 presents the frequency, by population category, with which systems have no treatment, one type of treatment, or multiple methods of treatment. Note that, in this table, both of the following cases would be counted as a system with two methods of treatment: (1) a system with two treatment plants, each providing a different type of treatment for a different entry point, and (2) a system with two methods of treatment for a single entry point.

[17] Furthermore, EPA's analysis of information from the AWWA Waterstat database for systems serving more than 50,000 people resulted in treatment frequencies similar to those here. The similarity of these results confirms that the sample used here is unbiased for system size categories where a comparison is possible.

Exhibit 6.3. Percent of Ground Water Systems with Treatment								
	Population Category							
Treatment Category	**Less than 100**	**101 to 500**	**501 to 1,000**	**1,001 to 3,300**	**3,301 to 10,000**	**10,001 to 50,000**	**50,001 to 100,000**	**More than 100,000**
Disinfection	52.8%	77.9%	84.0%	79.7%	86.8%	96.5%	86.3%	96.4%
Aeration	1.5%	6.3%	17.1%	19.9%	29.7%	33.0%	49.1%	44.1%
Oxidation	3.2%	6.6%	9.4%	4.2%	10.9%	9.3%	18.6%	5.4%
Ion Exchange	0.7%	1.6%	3.8%	1.9%	4.6%	3.3%	1.2%	0%
Reverse Osmosis	0%	1.2%	0%	0.9%	1.2%	0.7%	1.2%	0%
GAC	0%	0.5%	0%	0.4%	0%	6.7%	7.5%	9.0%
PAC	0%	0%	0%	0%	0.2%	0.3%	0%	1.8%
Filtration	11.8%	8.0%	15.9%	14.9%	29.5%	29.6%	50.3%	51.4%
Coagulation/ Flocculation	1.5%	5.4%	4.2%	3.4%	8.1%	15.1%	24.2%	25.2%
Lime/Soda Ash Softening	2.1%	3.7%	4.1%	5.2%	7.0%	12.2%	17.4%	32.4%
Recarbonation	0%	0.5%	0%	1.1%	3.0%	6.1%	7.5%	10.8%
Note: Percentages shown are weighted for item-level nonresponse.								

Exhibit 6.4. Percent of Surface Water Systems with Treatment								
	Population Category							
Treatment Category	**Less than 100**	**101 to 500**	**501 to 1,000**	**1,001 to 3,300**	**3,301 to 10,000**	**10,001 to 50,000**	**50,001 to 100,000**	**More than 100,000**
Disinfection	92.8%	94.1%	100.0%	100.0%	96.0%	98.0%	100.0%	100.0%
Aeration	0%	0%	1.4%	5.5%	8.5%	3.5%	10.3%	14.3%
Oxidation	0%	2.0%	7.2%	5.8%	7.7%	10.5%	5.7%	4.6%
Ion Exchange	0%	0%	0%	0%	0%	0%	0%	0%
Reverse Osmosis	0%	0%	0%	0%	0%	0%	0%	0%
GAC	3.9%	4.3%	1.4%	2.3%	4.7%	10.2%	14.9%	11.2%
PAC	0%	2.0%	3.0%	4.6%	18.6%	24.6%	34.2%	45.9%
Filtration	78.5%	71.2%	79.3%	81.7%	86.5%	96.3%	88.0%	93.4%
Coagulation/ Flocculation	27.5%	52.6%	70.2%	78.5%	95.4%	94.5%	93.7%	99.5%
Lime/Soda Ash Softening	3.9%	8.1%	20.5%	17.5%	10.8%	6.9%	5.7%	5.1%
Recarbonation	0%	0%	0%	0%	0%	0%	1.1%	5.1%
Note: Percentages shown are weighted for item-level nonresponse.								

Exhibit 6.5. Percent of Systems with Multiple Categories of Treatment								
Number of Treatment Categories	**Population Category**							
	Less than 100	**101 to 500**	**501 to 1,000**	**1,001 to 3,300**	**3,301 to 10,000**	**10,001 to 50,000**	**50,001 to 100,000**	**More than 100,000**
Ground Water Systems								
No treatment	42.5%	19.0%	16.0%	18.4%	13.1%	0.9%	11.2%	0%
1 treatment	43.9%	63.3%	57.4%	55.8%	45.1%	52.6%	22.4%	28.8%
2 treatments	12.3%	9.4%	7.0%	10.7%	8.3%	12.9%	14.9%	18.0%
3 treatments	0.4%	5.8%	12.3%	9.4%	19.6%	11.8%	11.2%	18.9%
4 treatments	1.0%	0.6%	6.6%	3.2%	9.4%	13.0%	25.5%	19.8%
5 treatments	0%	1.4%	0.8%	1.4%	4.2%	6.2%	13.7%	10.8%
6 treatments	0%	0.6%	0%	1.1%	0.3%	2.6%	1.2%	3.6%
Surface Water Systems								
No treatment	7.2%	5.9%	0%	0%	0%	2.0%	0%	0%
1 treatment	14.3%	23.0%	15.9%	7.5%	1.9%	0.7%	4.0%	0.5%
2 treatments	43.3%	17.5%	18.9%	14.0%	12.1%	60.6%	2.3%	4.6%
3 treatments	35.2%	40.3%	31.7%	60.8%	49.6%	22.6%	43.4%	35.7%
4 treatments	0%	11.4%	33.6%	10.5%	28.7%	13.6%	40.0%	40.3%
5 treatments	0%	2.0%	0%	7.3%	7.7%	13.6%	6.9%	12.2%
6 treatments	0%	0%	0%	0%	0%	0.5%	3.4%	6.6%
Note: Percentages shown are weighted for item-level nonresponse.								

Exhibit 6.6 presents similar data for groundwater systems at the individual entry point level. That is, the table describes the frequency with which an individual entry point is untreated or has a treatment plant providing one or more categories of treatment. Because the majority of surface water systems have only one entry point, percentages at the entry point level are nearly the same as those presented in Exhibit 6.5. Therefore, Exhibit 6.6 does not include data for surface water systems.

The data shown in Exhibits 6.3 through 6.6 suggest some patterns. Large systems appear to use more treatment methods and more advanced treatment methods than small systems. Also, surface water systems seem more likely to be treated than ground water systems. Additional evaluation of these data are necessary to incorporate these findings into analytical models.

Exhibit 6.6. Percent of Entry Points in Ground Water Systems with Multiple Methods of Treatment								
	Population Category							
Number of Treatement Categories	**Less than 100**	**101 to 500**	**501 to 1,000**	**1,001 to 3,300**	**3,301 to 10,000**	**10,001 to 50,000**	**50,001 to 100,000**	**More than 100,000**
No treatment	50.4%	24.8%	24.9%	26.9%	26.4%	7.0%	39.6%	17.4%
1 treatment	38.0%	61.3%	57.2%	56.2%	50.2%	70.1%	38.0%	52.1%
2 treatments	10.4%	7.8%	4.5%	7.2%	7.6%	11.4%	10.5%	18.6%
3 treatments	0.3%	4.3%	8.6%	6.6%	9.7%	5.7%	4.5%	6.4%
4 treatments	0.8%	0.4%	4.2%	1.8%	4.1%	3.9%	4.4%	1.6%
5 treatments	0%	1.0%	0.5%	0.8%	1.9%	1.2%	2.9%	3.9%
6 treatments	0%	0.4%	0%	0.6%	0.1%	0%	0.2%	0%
Note: Percentages shown are weighted for item-level nonresponse.								

7: Non-Community Water Systems

As discussed in Chapter 2, EPA's Safe Drinking Water Information System (SDWIS) database includes information on non-community water systems (NCWSs), as well as the community water systems (CWSs) that are the focus of previous chapters of this report. According to SDWIS, CWSs serve more than 90 percent of the total public water system population. Therefore, in previous drinking water regulations, EPA has used CWS flows to model NCWS flows. NCWS flows, however, are generally substantially lower than typical CWS flows. Furthermore, while NCWSs make up a small percentage of the population served, these systems actually comprise two-thirds of the total number of public water systems regulated under the SDWA.

The goal of this chapter is to provide an improved characterization of the NCWS universe. This quantitative discussion of NCWS model systems is based on data extracted from SDWIS in November 1997. Data limitations constrain the analysis of NCWSs. NCWSs are modeled separately from CWS because of inherent differences between the two types of systems and the lack of national NCWS survey data (i.e., comparable to the CWSS). The modeling approach presented herein for NCWSs uses SDWIS data and relies on various references for typical water consumption patterns for various types of NCWSs. Though not addressed here, system differences attributable to regional setting, variations in exposure routes related to system type, system residence times, water storage capabilities, and existing treatment profiles may be considered in subsequent efforts.

7.1 Overview of Non-Community Water System Population

SDWIS identifies 115,948 NCWSs in the United States (Source: 1997 SDWIS frozen database[18]). These NCWSs represent more than 67 percent of the total number of public water systems. Water sources for these systems include ground water, surface water (including ground water under the direct influence of surface water), and purchased water. Exhibit 7.1 presents a breakdown by water source of the number of NCWSs as reported in SDWIS. Based on SDWIS data, NCWSs serve more than 25-million people. NCWSs that serve less than 10,000 persons per system serve more than 15-million people of this total, with about 97 percent using ground water. As discussed below, systems that reported serving more than 10,000 people are treated separately in this analysis (see Appendix H).

Exhibit 7.1. Non-Community Water Systems by Water Source

Water Source	Number of Systems
Ground Water	112,214
Surface Water	2,119
Purchased Water	1,613
Other	2
Total	115,948

[18] The number of NCWSs reported in this chapter differ from those reported in Chapter 2, which reflect 1998 SDWIS data. Updated 1998 estimates were not available for use in the Chapter 7 analysis.

Exhibit 7.2 provides summary statistics for all NCWSs in SDWIS, sorted by water source and mean population served. Approximately 95 percent of the NCWSs serve less than 500 people, with about only 0.3 percent of NCWSs serving more than 3,300 persons a day. The remainder of this evaluation does not separate systems into population categories except to differentiate large NCWSs serving more than 10,000 persons a day. According to SDWIS, these large systems account for less than 0.1 percent of the total number of systems, but serve nearly one third of the total NCWS population (these systems are identified in Appendix H). Large NCWSs warrant separate evaluation because SDWIS data for some of these systems may be in error (e.g., systems reporting yearly population served rather than daily).

The remainder of this chapter focuses on the 112,214 ground water systems and 2,109 surface water systems serving fewer than 10,000 persons. Purchased water systems are excluded from discussion as these are effectively either CWS or NCWS customers.

Exhibit 7.2. Non-Community Water Systems by Population Range

Population Range	Ground Water Systems	Surface Water Systems	Total Systems
Less than 100	81,483	1,145	82,628
101-500	25,411	654	26,065
501-1,000	3,747	154	3,901
1,001-3,300	1,276	114	1,390
3,301-10,000	201	42	243
10,001-50,000	74	7	81
50,001-100,000	12	3	15
Greater than 100,000	10	0	10
Subtotal	112,214 (112,118)*	2,119 (2,109)*	114,333 (114,227)*

* Numbers in parentheses indicate values excluding all systems serving more than 10,000 persons.
Note: Totals shown exclude purchased water systems, systems using ground water under the direct influence of surface water, and systems reporting source water type as “other.”

A key characteristic for NCWSs is the distinction between transient and non-transient systems. The distinction is an important one, since regulations for chronic contaminants are not applied to transient water systems. A typical non-transient system may supply drinking water to employees (e.g., manufacturing facilities) or extended-stay residents (e.g., nursing homes), while a typical transient system may supply drinking water to service areas with short term and variable (i.e., transient) populations (e.g., amusement parks and restaurants). The SDWIS inventory on which this chapter is based initially categorized 95,858 systems as transient and 20,090 as non-transient (94,389 transient and 19,766 non-transient excluding purchased water systems, ground water systems under the direct influence of surface water, systems with source water specified as “other,” and systems serving more than 10,000 people).

The distinction between transient and non-transient systems can be unclear. For instance, SDWIS classifies churches as both transient and non-transient systems. While the population served can vary (i.e., it varies for certain days of the week and throughout the year), most churches serve the same individuals on a year-round basis, suggesting a classification as non-transient. However, well over 90 percent are classified as transient in SDWIS. Likewise, 10 percent of schools and 20 percent of daycare centers are classified as transient systems.

Given the “grey areas” between transient and non-transient systems, individual systems might be miscategorized in SDWIS. SDWIS ideally reflects the basis for distinguishing applicability of regulations; however, controversy over classifications could arise with regulations that require significant capital expenditures for compliance. Accordingly, this chapter presents the breakdown of transient versus non-transient systems based on the initial SDWIS inventory (Section 7.2) as well as a breakdown based on best professional judgment (Section 7.4). The analyst should carefully consider uses to be made of these data in determining which set to apply.

7.2 Service Area Classification and Population Served

SDWIS characterizes NCWSs by type of service and population served, among other variables.[19] In SDWIS, each NCWS is characterized by up to six service area type codes (see Exhibit 7.3). To develop a simple model of NCWSs, one service area type was assigned to each system (or multi-use systems were grouped, where appropriate). For purposes of the evaluation, these service area types are split into two categories: “specific” (those that narrowly define a population served, such as “daycare center”) and “general” (those that lack a usable designation, such as “other area”).

A review of SDWIS data found that approximately two-thirds of the NCWSs are codified as *general* service areas. A brief review of those systems suggests they are quite different from those with specific designations. To better define this segment of the NCWS population, an in depth evaluation of these facilities was performed. This multistep evaluation is described in Section 7.2.1.

Exhibit 7.3. SDWIS Service Area Classifications

***Specific* Service Area Classifications:**	***General* Service Area Classifications:**
- daycare center	- industrial/agricultural
- highway rest area	- institution
- hotel/motel	- no service area
- interstate carrier	- other area
- medical facility	- other non-transient area
- mobile home park	- other residential area
- restaurant	- other transient area
- school	- recreation area
- service station	- residential area
- summer camp	
- wholesaler	

[19] Characterizing the service area is important for NCWSs because service area directly impacts the exposure scenario. For instance, assumptions of daily water consumption for residents is inapplicable to restaurant customers.

7.2.1 General Approach

In general, modeling NCWSs requires categorizing all systems into *specific* service areas. Systems with one *specific* service area were simply categorized as presented in SDWIS. Systems with a *general* service area or more than one service area required more detailed examination. Specifically, systems with at lest one general service area were categorized into four distinct service area categories, depending on the service area types reported for each individual NCWS. A summary of the categorization process is presented in Exhibit 7.4.

Exhibit 7.4. Approach to NCWS Service Area Categorization

Category Number	Number of Specific Service Areas	Number of General Service Areas	Final Categorization Criteria
1	1	any number	Categorized solely on specific service area as reported in SDWIS (e.g., "day care center")
2	2 or more	any number	Given "Mixed known" category and subcategorized separately according to the specific service area reported in SDWIS; e.g., "mixed known with daycare center"
3	0	1	Recategorized using best professional judgement by name of system (process described in detail in Section 7.2.2)
4	0	2 or more	Given a "Mixed unknown" category for further categorization using best professional judgement by name of system (using the same process as category 3)

Categories 1 and 2 were sorted and represented by a specific service area. Categories 3 and 4 represent the body of NCWSs reporting only general service areas. Further analysis to sort these systems by specific service area is summarized in the next section.

7.2.2 Characterization of General Service Area Classifications

To further characterize the makeup of the NCWSs reporting only general service areas, a random sample of NCWSs was collected from each general service area classification. The sample size was based on achieving a 95 percent probability that no service area representing more than 0.5 percent of the NCWS population would be missed in the process. As depicted in Exhibit 7.5, a total of 1,152 NCWSs were selected from the universe of 76,179 NCWSs identified with general service area types. The systems were further subcategorized according to their initial categorization in SDWIS as "transient" or "non-transient."

Exhibit 7.5. Random Sampling of Non-Community Water Systems with General Service Areas

General Service Area Classification	Transient*		Non-Transient*	
	Sampling Frame	Sample Size	Sampling Frame	Sample Size
Industrial/Agricultural	1,166	9	4,428	226
Institution	492	9	374	14
No Service Area	22,000	200	3,650	180
Other Area	3,539	26	552	21
Other Non-Transient Area	687	9	1,802	80
Other Residential Area	274	9	793	7
Other Transient Area	17,939	161	627	21
Recreation Area	16,383	141	221	7
Residential Area	631	9	123	7
Mixed Unknown	1,027	9	171	7
Total	64,138	582	12,041	570

* As initially categorized in SDWIS. See Section 7.4 for further discussion of transient versus non-transient categorization.

Specific to any Delphi approach, the goal of this effort was to make forecasts that systematically use insights and assessments of selected specialists. A service area type was assigned to each of the 1,152 general NCWSs using the consensus of best professional judgment from four reviewers based on the system name as provided in SDWIS. A total of 59 service area types were identified as a result of this process, including the 11 SDWIS specific types. A list of the 59 service area types is provided in Exhibit 7.6

The sampled NCWSs were scaled-up based on the ratio of the number of systems in the sampling frame to the number of systems in the random sample. For example, facilities coded in the “transient” and “other area” category were scaled up using a 3,539/26 scaling factor.

Exhibit 7.6 Non-community Water System Service Areas

Existing SDWIS Specific Categories

Code	Service Area Type [SIC Code]
(DC)	Daycare Center [8351]
(HRA)	Highway Rest Area [NO SIC]
(HM)	Hotel/Motel [70] (includes rooming and boarding houses, lodges, and resorts)
(IC)	Interstate Carrier (includes truck stops, bus and railroad terminals, airports, couriers, postal service)
(MF)	Medical Facility [80]
(MHP)	Mobile Home Park [6515]
(R)	Restaurant [581]
(S)	School [82] (includes colleges, vocational schools, dance studios, and universities or places of higher learning)
(SS)	Service Station [5541 and 75]
(SC)	Summer Camp (include basketball camps, baseball camps etc.) [7032]
(WPP)	Water Wholesaler or Producer (include washeterias)

Categories Identified in Sampling of General Service Areas

Code	Service Area Type [SIC Code]
(AG)	Agricultural [01, 02, and 07]
(AP)	Air Park [4581]
(B)	Bowling Centers [7933]
(C)	Construction [15, 16, and 17]
(CH)	Churches [866]
(CRV)	Campground or RV Parks [7033]
(FD)	Fire Departments [9224]
(FP)	Federal Parks [9512]
(FS)	Forest Service [9512]
(GC)	Golf and Country Clubs [7992]
(L)	Laundries, Including Industrial Laundries [721]
(LIB)	Libraries [8231]
(LFL)	Landfill [4953]
(M)	Mining [10, 12, 13, and 14]
(MAMU)	Amusement Parks (includes Fairgrounds and Water Parks) [7996]
(MB)	Military Bases [9711]
(MFCC)	Industrial and Commercial Machinery and Computer Equipment. [35]
(MFCE)	Electronic and Electrical Equipment and Components, Except Computers [36]
(MFCH)	Chemicals and Allied Products [28]
(MFF)	Furniture and Fixtures [25]
(MFI)	Miscellaneous Mfg. Industries [39]
(MFLL)	Leather and Leather Products [31]
(MFM)	Fabricated Metal Products, Not Transportation. [34]
(MFO)	Food and Kindred Products [20]
(MFP)	Paper and Allied Products [26]
(MFPE)	Petroleum Refining and Related Industries [29]
(MFPM)	Primary Metal Industries [33]
(MFPR)	Printing, Publishing, and Allied Industries [27]
(MFRU)	Rubber and Miscellaneous Plastics Products [30]
(MFSC)	Stone, Clay, Glass, and Concrete Products [32]
(MFT)	Tobacco Products [21]
(MFTE)	Transportation Equipment [37]
(MFTX)	Textile Mill Products [22]
(MFTX)	Apparel and Other Finished Products [23]
(MFW)	Lumber and Wood Products, except furniture [24]
(MFWW)	Measuring, Analyzing, and Controlling Instruments, Photo, Medical and Optical Goods; Watches and Clocks [38]
(MLC)	Migrant Labor Camps [0761]
(MREC)	Miscellaneous Recreation Services [799] (excluding amusement parks)
(MU)	Museums [84]
(NH)	Nursing Homes [805]
(OP)	Office Parks [6512]
(PRI)	Prisons [9223]
(RCC)	Racing, including track operation [7948]
(RET)	Retailers (Non-food related) [53 and 55]
(RETF)	Retailers (Grocery Stores, Fruit/Vegetable Markets, Meat and Fish Markets, Dairy Products, Bakeries, etc.) [54]
(SP)	State Parks [9512]
(UT)	Non-Water Utilities (includes power plants, natural gas, electric companies) [491,492]
(ZG)	Zoological Gardens [84] (e.g., arboretums)

7.2.3 Results

Based on the categorization approach presented in Sections 7.2.1 and 7.2.2, combined estimates were developed for the number of NCWSs, the average population served, and the total population served for each identified service classification. Exhibit 7.7 presents these results. Exhibit 7.8 identifies the estimated population served for each service classification (e.g., for restaurants, the average population served is represented as the number of customers daily).

The totals derived after application of the categorization approach are slightly different than those presented previously in Exhibit 7.2. These differences, however, are relatively insignificant (114,227 systems originally identified in SDWIS versus 114,726 systems after categorization, or a 0.44 percent difference). These differences are a result of the approach used to characterize the systems with general service classifications and subsequent rounding.[20]

This evaluation does not summarize NCWSs by population category since, as noted in Section 7.1, more than 95 percent of all NCWSs serve fewer than 500 persons a day, with nearly 99 percent serving 1,000 persons or fewer. From a regulatory impact standpoint, additional stratifications would have little impact on the accuracy of models. Rather, the diversity of ownership is the primary variable of interest.

A separate evaluation of large systems (i.e., those serving more than 10,000 persons) was performed to identify the types of systems represented. These systems were categorized based on system name, similar to the approach used for identifying the smaller systems. Specifically, each of the 106 systems was assigned a service classification based on best professional judgment. A list of these largest NCWSs, with the assigned code, is provided as Appendix H. Of the systems serving more than 10,000 persons, approximately two-thirds are State parks, with highway rest areas, miscellaneous amusement parks, and campgrounds accounting for most of the rest. Many of the systems reporting a daily population served of greater than 10,000 appear to be reporting errors. Many of the populations appear to be monthly, yearly, or peak daily figures. For example, campgrounds reporting populations of 12,000 persons or more and highway rest areas serving over 60,000 people per day. Based on a best-professional-judgment evaluation of the NCWSs reporting service populations greater than 10,000, approximately two-thirds appear to be incorrectly recorded. As such, the percentage of systems serving greater than 10,000 is likely even smaller than presented here.

[20] For the same reasons, similar small differences will exist between the population served totals shown here and similar figures derived directly from SDWIS (see Appendix A).

Exhibit 7.7. Non-Community Water System Population Served by Service Area Type*

Service Area Type	Transient			Non-Transient			Total		
	Number of Systems	Average Population	Total Population	Number of Systems	Average Population	Total Population	Number of Systems	Average Population	Total Population
Daycare Centers	200	51	10,213	809	76	61,653	1,009	71	71,866
Highway Rest Areas	1,311	394	516,369	15	407	6,105	1,326	394	522,474
Hotels/Motels	8,312	67	558,443	351	133	46,680	8,663	70	605,123
Interstate Carriers	128	88	11,257	287	123	35,221	415	112	46,478
Medical Facilities	672	310	208,623	367	393	144,061	1,039	339	352,684
Mobile Home Parks	1,220	55	66,797	104	185	19,236	1,324	65	86,033
Restaurants	25,422	89	2,255,959	418	370	154,528	25,840	93	2,410,487
Schools	923	163	150,365	8,414	358	3,015,155	9,337	339	3,165,520
Service Stations	3,123	105	326,644	53	230	12,177	3,176	107	338,821
Summer Camps	6,149	125	765,742	46	146	6,711	6,195	125	772,453
Water Wholesalers	1,430	553	791,429	266	173	46,075	1,696	494	837,504
Agricultural Products/Services	690	33	22,770	368	76	27,968	1,058	48	50,738
Airparks	476	141	67,116	101	60	6,060	577	127	73,176
Bowling Centers	331	70	23,170	0	0	0	331	70	23,170
Construction	0	0	0	99	53	5,247	99	53	5,247
Churches	11,621	112	1,301,552	230	50	11,500	11,851	111	1,313,052
Campgrounds/RV Parks	5,568	115	639,160	123	160	19,680	5,691	116	658,840
Fire Departments	331	38	12,578	41	98	4,018	372	45	16,596
Federal Parks	655	143	93,665	20	39	780	675	140	94,445
Forest Service	752	50	37,600	107	42	4,494	859	49	42,094
Golf and Country Clubs	2,352	108	254,016	116	101	11,716	2,468	108	265,732
Landfills	0	0	0	78	44	3,432	78	44	3,432
Libraries	111	30	3,330	0	0	0	111	30	3,330
Mining	0	0	0	119	113	13,447	119	113	13,447
Amusement Parks	603	146	88,038	159	418	66,462	762	203	154,500
Military Bases	116	25	2,900	95	395	37,525	211	192	40,425
Migrant Labor Camps	620	45	27,900	33	63	2,079	653	46	29,979
Misc Recreation Services	3,512	96	337,152	259	87	22,533	3,771	95	359,685
Museums	252	140	35,280	0	0	0	252	140	35,280
Nursing Homes	0	0	0	130	107	13,910	130	107	13,910
Office Parks	1,976	100	197,600	950	136	129,542	2,926	112	327,142

Exhibit 7.7. Non-Community Water System Population Served by Service Area Type* (Continued)

Service Area Type	Transient			Non-Transient			Total		
	Number of Systems	Average Population	Total Population	Number of Systems	Average Population	Total Population	Number of Systems	Average Population	Total Population
Prisons	0	0	0	67	1,820	121,940	67	1,820	121,940
Racing, including Track Operations	116	500	58,000	0	0	0	116	500	58,000
Retailers (Non-Food Related)	3,288	56	184,128	695	174	120,775	3,983	77	304,903
Retailers (Food Related)	2,707	53	142,988	142	322	45,724	2,849	66	188,712
State Parks	6,329	133	842,518	83	165	13,712	6,412	134	856,230
Non-Water Utilities	241	25	6,025	497	170	84,621	738	123	90,646
Zoological Gardens	110	30	3,300	0	0	0	110	30	3,300
Manufacturing: Food	533	297	158,301	768	372	285,910	1,301	341	444,211
Manufacturing: Machinery	0	0	0	20	2,000	40,000	20	2,000	40,000
Manufacturing: Electrical	0	0	0	79	27	2,133	79	27	2,133
Manufacturing: Chemicals	0	0	0	138	90	12,384	138	90	12,384
Manufacturing: Furniture	110	25	2,750	46	32	1,472	156	27	4,222
Manufacturing: Miscellaneous	333	27	8,991	1,760	156	275,146	2,093	136	284,137
Manufacturing: Fabricated Metal	110	30	3,300	466	94	43,804	576	82	47,104
Manufacturing: Paper & Allied	0	0	0	160	241	38,560	160	241	38,560
Manufacturing: Petroleum Refining	0	0	0	105	343	35,855	105	343	35,855
Manufacturing: Primary Metals	0	0	0	79	333	26,278	79	333	26,278
Manufacturing: Printing	0	0	0	20	200	4,000	20	200	4,000
Manufacturing: Rubber & Plastics	0	0	0	40	50	2,000	40	50	2,000
Manufacturing: Stone, Clay, Glass	111	25	2,775	118	247	29,146	229	139	31,921
Manufacturing: Tobacco Products	0	0	0	20	75	1,500	20	75	1,500
Manufacturing: Transportation	0	0	0	40	27	1,080	40	27	1,080
Manufacturing: Textiles	0	0	0	85	407	34,590	85	407	34,590
Manufacturing: Lumber & Wood	111	25	2,775	180	85	15,300	291	62	18,075
Unknown Service Areas	220	25	5,500	43	392	16,856	263	85	22,356
Mixed Service Areas	1,558	138	214,345	184	504	92,797	1,742	176	307,142
Total	94,733	110	10,441,364	19,993	264	5,273,578	114,726	137	15,714,942

* Covers systems serving less than 10,000 people per system. Appendix H describes systems serving more than 10,000 people per system.

Exhibit 7.8. Explanation of Population Served for Non-Community Water Systems

Service Area Type	Population Served Represents
Daycare Centers	Daily occupancy and employees
Highway Rest Areas	Daily visitors
Hotels/Motels	Daily occupancy and employees
Interstate Carriers	Employees and/or daily passengers
Medical Facilities	Patients and employees
Mobile Home Parks	Daily residents
Restaurants	Daily customers and employees
Schools	Students and employees
Service Stations	Daily customers
Summer Camps	Daily campers
Water Wholesalers	Daily customers
Agricultural Products/Services	Employees
Airparks	Daily visitors and employees
Bowling Centers	Daily customers and employees
Construction	Daily workers
Churches	Average congregation
Campgrounds/RV Parks	Daily visitors
Fire Department	Population protected
Federal Parks	Daily visitors
Forest Service	Daily visitors and/or employees
Golf and Country Clubs	Daily patrons and employees
Landfills	Employees
Libraries	Employees
Mining	Employees
Amusement Parks	Daily visitors and employees
Military Bases	Personnel
Migrant Labor Camps	Daily occupancy
Miscellaneous Recreation Services	Daily visitors and employees
Museums	Daily visitors and employees
Nursing Homes	Occupants and employees
Office Parks	Employees
Prisons	Inmates and employees
Racing, including Track Operations	Daily visitors and employees
Retailers (Non-Food Related)	Daily customers and employees
Retailers (Food Related)	Daily customers and employees
State Parks	Daily visitors
Non-Water Utilities	Employees
Zoological Gardens	Daily visitors and employees
Manufacturing	Employees
Unknown Service Areas	Unknown
Mixed Service Areas	Depends on types represented

7.3 Average and Design Flows

Estimates of the average water-use rate, in gallons per person per day for each of the 59 service classifications, were developed using a variety of literature sources. Where water-use rates were not identified in the literature for a given service classification, best professional judgment was used to estimate a usage for similar facilities. Exhibit 7.9 provides a summary of these average water-use rates, including the basis for best professional judgment determinations. Based on these average water consumption rates and the estimated population served for each service area type shown in Exhibit 7.7, average daily system flows within each service classification were estimated. Exhibit 7.10 provides these results. Similar to CWSs, average daily flows are an important input in estimating operation and maintenance costs for regulatory analysis purposes.

Design flows for NCWS are also an important input in estimating capital costs for regulatory analysis purposes. Design flows for NCWSs can be estimated based on design-to-average flow ratios. Design-to-average flow ratios for NCWSs may differ from those for CWSs. Some NCWS may have greater ratios than CWSs because storage is typically not available and systems are designed to accommodate larger variations in demand. However, some NCWSs may have lower design-to-average ratios because their demand changes little from day to day. In general the design-to-average flow ratios for ground water CWSs are thought to be a reasonable approximation of ratios for NCWSs of similar size.

The following approach was used to estimate design flows for NCWSs that have similar design-to-average flow ratios to ground water CWSs:

(1) Use the average daily flow per system for each NCWS service area type (Exhibit 7.10) to back-calculate an equivalent, or “virtual” population using the CWS average daily flow equation for public ground water systems (Chapter 4).

(2) Use the equivalent, or “virtual” population produced in step 1 to estimate the design flow for each NCWS service are type using the CWS design flow regression equation for public ground water systems (Chapter 4).

The results of these calculations are presented in Exhibit 7.10

Exhibit 7.9. Average Water Use Assumptions for Non-Community Water Systems (gallons per person per day)

Specific Service Area Type	**Water Use**	**Source with "assumption"**
Daycare Centers	15	(1) *"Day camp"*
Highway Rest Areas	5	(1)
Hotels/Motels	65	(1)
Interstate Carriers	5	(1)
Medical Facilities	100	Best Professional Judgment
Mobile Home Parks	100	(4)
Restaurants	8.5	(3)
Schools	25	(3)
Service Stations	10	(1)
Summer Camps	42.5	(1)
Water Wholesalers	100	(4)

General Service Area Type	**Water Use**	**Source with "assumption"**
Agricultural Products/Services	100	Best Professional Judgment
Airparks	4	(3)
Bowling Centers	3	(1) *"Movie theater"*
Construction	3	Best Professional Judgment
Churches	10	(1) *"Picnic with Toilet Facilities"*
Campgrounds/RV Parks	45	(1)
Fire Departments	100	(4)
Federal Parks	10	(3)*"Picnic with Toilet Facilities"*
Forest Service	5	(1) *"Campsite no toilet, bath, or shower"*
Golf/Country Clubs	25	(3)
Landfills	25	(1) *"Day workers"*
Libraries	15	(2)
Migrant Labor Camps	50	(3) *"Construction workers"*
Military Bases	100	(4) *"Residential"*
Mines	25	(1) *"Day worker"*
Miscellaneous Amusement Parks	20	(3) *"Picnic with toilet, shower, etc."*
Miscellaneous Recreation Areas	5	(3) *"Theater"*
Museums	10	(2) *"Department store"*
Nursing Homes	100	(4) *"Residential"*
Office Parks	15	(2) *"Office"*
Prisons and jails	100	(1) *"Institution"*
Race Tracks	5	(1) *"Fairgrounds"*
Retailers (excluding food)	10	(2) *"Department store"*
Retailers (food)	8.5	(3) *"Restaurant"*
State Parks	7.5	(1) *"Picnic with toilet facility"*
Utilities	25	(1) *"Day workers"*
Zoological Gardens	25	(1) *"Day workers"*
Manufacturing (Food and Kindred Products)	35	Best Professional Judgment
All Other Manufacturing Categories	25	(1)

Sources:

1 Salvato, Joseph A. *Environmental Engineering and Sanitation*, 4th Edition
2 Metcalf & Eddy. 1991. *Wastewater Engineering Treatment, Disposal, and Reuse.*
3 Corbitt, Robert A. 1990. *Standard Handbook of Environmental Engineering.*
4 Community Water System Survey (CWSS), 1997.

Exhibit 7.10 Non-community Water System Flow Rates by Service Area Type

Service Area Type	Average Flow per Capita per Day (gpd)	Average Daily Flow per System (gpd)	Design Flow (gpd)	Design/Average Ratio
Daycare Centers	15	1,068	5,379	5.0
Highway Rest Areas	5	1,970	9,651	4.9
Hotels/Motels	65	4,540	21,428	4.7
Interstate Carriers	5	560	2,902	5.2
Medical Facilities	100	33,945	146,450	4.3
Mobile Home Parks	100	6,498	30,181	4.6
Restaurants	8.5	793	4,045	5.1
Schools	25	8,476	38,903	4.6
Service Stations	10	1,067	5,371	5.0
Summer Camps	42.5	5,299	24,839	4.7
Water Wholesalers	100	49,381	209,516	4.2
Agricultural Products/Services	100	4,796	22,578	4.7
Airparks	4	507	2,640	5.2
Bowling Centers	3	210	1,137	5.4
Construction	3	159	871	5.5
Churches	10	1,108	5,569	5.0
Campgrounds/RV Parks	45	5,210	24,437	4.7
Fire Departments	100	4,461	21,072	4.7
Federal Parks	10	1,399	6,960	5.0
Forest Service	5	245	1,317	5.4
Golf and Country Clubs	25	2,692	13,004	4.8
Landfills	25	1,100	5,531	5.0
Libraries	15	450	2,355	5.2
Mines	25	2,825	13,618	4.8
Miscellaneous Amusement Parks	20	4,055	19,235	4.7
Military Bases	100	19,159	84,795	4.4
Migrant Labor Camps	50	2,295	11,168	4.9
Miscellaneous Recreation Areas	5	477	2,489	5.2
Museums	10	1,400	6,964	5.0
Nursing Homes	100	10,700	48,604	4.5
Office Parks	15	1,677	8,275	4.9
Prisons	120	218,400	867,156	4.0
Racing, including Track Operations	5	2,500	12,117	4.8
Retailers (Non-Food Related)	10	766	3,912	5.1
Retailers (Food Related)	18.5	1,225	6,131	5.0
State Parks	7.5	1,002	5,057	5.0
Zoological Gardens	25	750	3,836	5.1
Manufacturing: Food	35	11,950	54,017	4.5
Manufacturing: Machinery	25	50,000	212,023	4.2
Manufacturing: Electronic Equipment	25	675	3,469	5.1
Manufacturing: Chemicals	25	2,243	10,927	4.9
Manufacturing: Furniture &	25	677	3,476	5.1

Service Area Type	Average Flow per Capita per Day (gpd)	Average Daily Flow per System (gpd)	Design Flow (gpd)	Design/Average Ratio
Manufacturing: Miscellaneous	25	3,394	16,227	4.8
Manufacturing: Fabricated Metal	25	2,044	9,999	4.9
Manufacturing: Paper & Allied	25	6,025	28,079	4.7
Manufacturing: Petroleum Refining	25	8,575	39,338	4.6
Manufacturing: Primary Metals	25	8,316	38,202	4.6
Manufacturing: Printing	25	5,000	23,497	4.7
Manufacturing: Rubber & Plastics	25	1,250	6,249	5.0
Manufacturing: Stone, Clay, Glass, etc	25	3,485	16,642	4.8
Manufacturing: Tobacco Products	25	1,875	9,205	4.9
Manufacturing: Transportation Equip.	25	675	3,469	5.1
Manufacturing: Textiles	25	10,174	46,317	4.6
Manufacturing: Lumber & Wood	25	1,553	7,688	5.0
Unknowns	25	2,125	10,375	5.0

7.4 Transient Versus Non-Transient Systems

As discussed earlier in this chapter, some question exists regarding the accuracy of the SDWIS subcategorizations of NCWS. The sampling procedure described in Section 7.2 could exacerbate such miscategorizations. For example, restaurants are considered transient systems. If, during the service classification sampling, one restaurant was selected that was miscategorized as non-transient, this would lead to a final estimate reflecting a much larger number of non-transient restaurants. In fact, the final estimate resulting from the service area sampling described in Section 7.2 included a number of systems that appeared to be miscategorized (e.g., it included some non-transient restaurants, transient manufacturing facilities, etc.).

While the categorization may be technically correct (e.g., restaurant could have more than 25 employees), an alternative breakdown of transient versus non-transient systems based on service class is offered to illustrate the potential difference between the existing SDWIS classifications and reality. Service area types (e.g., restaurants, service stations) whose populations are variable (e.g., representing customers, visitors, or guests) were classified as transient. Service classes (e.g., schools, manufacturing facilities) whose populations are consistent (e.g., representing employees or residents) were classified as non-transient. Some systems reasonably could be either transient or non-transient. The breakdown of transient versus non-transient for these systems was not changed. These system types included the following:

- Interstate Carriers: includes truck stops and bus and railroad terminals where the primary water users would be transient (e.g., passengers), but also includes freight depots and postal service operations where the primary water users would be employees (non-transient)

- Hotels: usually transient, but includes boarding houses in which the population might more appropriately be categorized as non-transient

- Medical Facilities: includes some extended stay facilities (e.g., nursing homes) that are non-transient

- Mobile Home Parks: includes some with seasonal populations (transient) and some that are more similar to CWSs (non-transient)

- Agricultural Products and Services: includes facilities more similar to retail food operations (transient) and facilities more similar to farms or food manufacturers (non-transient)

- Airparks: similar to interstate carriers

- Forest Service: includes areas that are primarily recreational (transient) and areas in lumber production where the primary users would be employees (non-transient)

Exhibit 7.11 summarizes the revised estimate of transient versus non-transient systems. This estimate makes the breakdown of transient versus non-transient systems consistent with types of service classes estimated for the population of NCWSs.

Exhibit 7.11. Alternative Classification of Transient versus Non-Transient Non-Community Water Systems

Service Area Type	Transient			Non-Transient			Total		
	Number of Systems	Average Population	Total Population	Number of Systems	Average Population	Total Population	Number of Systems	Average Population	Total Population
TRANSIENT SERVICE AREA TYPES									
Highway Rest Areas	1,326	394	522,474	0	0	0	1,326	394	522,474
Restaurants	25,840	93	2,410,487	0	0	0	25,840	93	2,410,487
Service Stations	3,176	107	338,821	0	0	0	3,176	107	338,821
Summer Camps	6,195	125	772,453	0	0	0	6,195	125	772,453
Water Wholesalers	1,696	494	837,504	0	0	0	1,696	494	837,504
Bowling Centers	331	70	23,170	0	0	0	331	70	23,170
Campgrounds/RV Parks	5,691	116	658,840	0	0	0	5,691	116	658,840
Federal Parks	675	140	94,445	0	0	0	675	140	94,445
Golf and Country Clubs	2,468	108	265,732	0	0	0	2,468	108	265,732
Libraries	111	30	3,330	0	0	0	111	30	3,330
Amusement Parks	762	203	154,500	0	0	0	762	203	154,500
Migrant Labor Camps	653	46	29,979	0	0	0	653	46	29,979
Misc. Recreation Services	3,771	91	359,685	0	0	0	3,771	91	359,685
Museums	252	140	35,280	0	0	0	252	140	35,280
Racing, including Track Operations	116	500	58,000	0	0	0	116	500	58,000
Retailers (Non-Food Related)	3,983	77	304,903	0	0	0	3,983	77	304,903
Retailers (Food Related)	2,849	66	188,712	0	0	0	2,849	66	188,712
State Parks	6,412	134	856,230	0	0	0	6,412	134	856,230
Zoological Gardens	110	30	3,300	0	0	0	110	30	3,300

Exhibit 7.11. Alternative Classification of Transient versus Non-Transient Non-Community Water Systems* (Continued)

Service Area Type	Transient			Non-Transient			Total		
	Number of Systems	Average Population	Total Population	Number of Systems	Average Population	Total Population	Number of Systems	Average Population	Total Population
NON-TRANSIENT SERVICE AREA TYPES									
Daycare Centers	0	0	0	1,009	71	71,866	1,009	71	71,866
Schools	0	0	0	9,337	339	3,165,520	9,337	339	3,165,520
Construction	0	0	0	99	53	5,247	99	53	5,247
Churches	0	0	0	11,851	111	1,313,052	11,851	111	1,313,052
Fire Departments	0	0	0	372	45	16,596	372	45	16,596
Landfills	0	0	0	78	44	3,432	78	44	3,432
Mining	0	0	0	119	113	13,447	119	113	13,447
Military Bases	0	0	0	211	192	40,425	211	192	40,425
Manufacturing (all categories)	0	0	0	5,432	192	1,028,050	5,432	192	1,028,050
Nursing Homes	0	0	0	130	107	13,910	130	107	13,910
Office Parks	0	0	0	2,926	112	327,142	2,926	112	327,142
Prisons	0	0	0	67	1,820	121,940	67	1,820	121,940
Non-Water Utilities	0	0	0	738	123	90,646	738	123	90,646
TRANSIENT AND NON-TRANSIENT SERVICE AREA TYPES									
Hotels/Motels	8,312	67	558,443	351	130	46,680	8,663	70	605,123
Interstate Carriers	128	88	11,257	287	126	35,221	415	112	46,478
Medical Facilities	672	310	208,623	367	393	144,061	1,039	339	352,684
Mobile Home Parks	1,220	55	66,797	104	185	19,236	1,324	65	86,033
Agricultural Products/Services	690	33	22,770	368	76	27,968	1,058	48	50,738
Airparks	476	141	67,116	101	60	6,060	577	127	73,176
Forest Service	752	50	37,600	107	42	4,494	859	49	42,094
Unknown Service Areas	220	25	5,500	43	292	16,856	263	85	22,356
Mixed Service Areas	1,558	138	214,345	184	504	92,797	1,742	176	307,142
	14,028	113	9,110,296	34,281	192	6,604,646	48,309	325	15,714,942

* Covers systems serving less than 10,000 people per system. Appendix H describes systems serving more than 10,000 people per system.

8: References

58 FR 51735. Executive Order 12866. "Regulatory Planning and Review." October 4, 1993.

Aschengrau, A, S. Zierler and A. Cohen. 1993. "Quality of Community Drinking Water and the Occurrence of Late Adverse Pregnancy Outcomes." Arch. Environ Health. 48:105-113.

Bauman, Duane D., et. al. Urban Water Demand Management and Planning, McGraw-Hill, New York, NY, 1998.

Corbitt, Robert A. Standard Handbook of Environmental Engineering. McGraw-Hill, New York, NY, 1990.

Cummins, Michael D. 1987. "Analysis of Flow Data." Report prepared for EPA, ODW, October 5th 199X

Dysard, J.A. 1999. "Trends in Privatization." Journal of the American Water Works Association. 91:44-47.

Linsley et. al. Water-Resources Engineering, 4th edition. Irwin McGraw-Hill, Inc., Boston, MA. 1992.

Salvato, Joseph A. Environmental Engineering and Sanitation, 4th Edition. John Wiley & Sons, New York, NY, 1992.

Tchobanoglous and Burton. Wastewater Engineering: Treatment, Disposal, and Reuse, 3rd Edition. McGraw-Hill, Inc., New York, NY.1991.

U.S. Environmental Protection Agency. Community Water System Survey. Office of Science and Technology, Office of Water, Washington, D.C., 1997.

Viessman and Hammer. Water Supply and Pollution Control, 6th edition. Addison-Wesley Longman, Inc., Menlo Park, California. 1998.

Appendix A: Population Served for Public Water Systems

Appendix B: Methodology for Quantifying the Cost Bias Caused by Retail Population Categorization of Systems

Appendix C: 1995 Community Water Systems Survey Questionnaire

Appendix D: Weighted Regression Methodology

Appendix E: Statistical Testing of Regressions for Different Categories of CWS

Appendix F: Bootstrapping Results for Ground Water System Entry Points

Appendix G: Distribution of Flow Among Entry Points

Appendix H: Non-Community Water Systems Serving Greater than 10,000 People